DOCUMENT RESUME

ED 391 827 TM 024 410

AUTHOR Raupp, Magdala; Kolb, Frances Arick, Ed.
TITLE Evaluation Management Handbook.
INSTITUTION NETWORK, Inc., Andover, MA.
PUB DATE Dec 90
NOTE 82p.
PUB TYPE Guides - Non-Classroom Use (055)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *Advocacy; Cost Estimates; Data Analysis; Data Collection; *Evaluation Methods; Evaluation Needs; Evaluation Problems; Evaluation Utilization; Evaluators; Program Development; Program Evaluation; Research Design; *Strategic Planning; *User Needs (Information)

ABSTRACT

This handbook is designed to help educators understand program evaluation processes so that they will be able to carry out evaluation tasks and talk knowledgeably to evaluators about what they want to find out from a program evaluation. The book reviews the purposes and uses of evaluation to help the reader identify the scope and direction of the evaluation process. Section 3 introduces the strategic steps in an evaluation and gives practical suggestions to avoid pitfalls. The 10 basic steps to program evaluation are delineated: (1) identify program goals; (2) state objectives; (3) decide on the scope of the evaluation; (4) formulate evaluation questions; (5) design the evaluation; (6) choose the method of information collection; (7) list evaluation activities and establish a timeline; (8) estimate the cost of the evaluation; (9) collect, analyze, and interpret the data; and (10) communicate and act on the evaluation findings. How to report and use findings is reviewed, and a glossary of evaluation terms is included. (Contains 26 references.) (SLD)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED 391 827

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
M. RAUPP
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

4981

# EVALUATION management HANDBOOK

DECEMBER 1990

The NETWORK, Inc.
300 BRICKSTONE SQUARE, SUITE 900, ANDOVER, MASSACHUSETTS 01810

2 BEST COPY AVAILABLE

# EVALUATION management HANDBOOK

December 1990

Produced by

Magdala Raupp
Author

Frances Arick Kolb
Editor

The NETWORK, Inc.
300 Brickstone Square, Suite 900
Andover, Massachusetts 01810

For further information, contact:

Publications Department
The NETWORK, Inc.
300 Brickstone Square, Suite 900
Andover, Massachusetts 01810
508-470-1080

© The NETWORK, Inc.
December 1990

# table of contents


| | | |
|---|---|---|
| Introduction | | 5 |
| Section 1: | QUESTIONS AND ANSWERS ABOUT THE HANDBOOK | 6 |
| Section 2: | IDEAS ABOUT EVALUATION: WHY? WHAT? HOW? | 8 |
| Section 3: | TEN BASIC STEPS TO EVALUATE YOUR PROGRAM | 11 |
| STEP #1: | Formulate Goals | 12 |
| STEP #2: | State Objectives | 14 |
| STEP #3: | Decide on the Scope of the Evaluation | 19 |
| STEP #4: | State the Evaluation Questions | 21 |
| STEP #5: | Design the Evaluation | 23 |
| STEP #6: | Choose the Methods of Information Collection | 28 |
| STEP #7: | List Evaluation Activities and Establish a Timeline | 47 |
| STEP #8: | Estimate the Cost of the Evaluation | 52 |
| STEP #9: | Collect, Analyze, and Interpret the Data | 55 |
| STEP #10: | Communicate and Act on the Evaluation Findings (Evaluation Report) | 58 |
| Section 4: | USING THE EVALUATION FINDINGS | 65 |
| Section 5: | SOME COMMON MISCONCEPTIONS ABOUT EVALUATION | 69 |
| Glossary | | 71 |
| Consumer Tips on Evaluation Services | | 75 |
| References | | 77 |

# introduction

The *Evaluation Management Handbook* is designed for people who: (1) have decided to carry out some of their program evaluation tasks themselves; (2) plan to contract their evaluation with an outside evaluator but want to be knowledgeable about the evaluation process; (3) need help in the creation of an evaluation plan for a new proposal; or, (4) want to use evaluation methodology to inform their work.

All project staff, not just a professional evaluator, should be a part of the evaluation process. Recent research has shown that the usefulness of any evaluation increases when

- the evaluators and the persons who will use the data work together as a team;
- professional leaders have some training in the evaluation process; and
- decision makers plan how they will use the evaluation results prior to the collection of the data.

Evaluation provides two important ingredients for a project's success. First, ongoing evaluation of a project ensures adherence to the goals and objectives of the project and to effective implementation. Through evaluation, the project staff will see areas where problems are occurring and make provisions for program improvement. Secondly, the evaluation will show whether the program has had any impact on its intended audience.

As always at The NETWORK, Inc., many support staff were of major importance to this publication. Special thank yous go to Clif Lund-Rollins, who helped produce the final manuscript, and to Nick Thorkelson, who created the book design and graphics.

# SECTION 1

## questions and answers about the handbook

### What is its purpose?

The purpose of the *Evaluation Management Handbook* is to help educators to understand program evaluation processes so that they will be able to carry out evaluation tasks and talk knowledgeably to an evaluator about what they want to find out from the evaluation of their program.

The handbook addresses questions such as: How can you use evaluation methods to improve program effectiveness? What evaluation instruments will you use? How good is the evaluation design? How will you report the results? How can you use the results of the evaluation? How good was the evaluation and how can it be improved?

The design of the handbook is based on recommendations from recent research on evaluation usefulness, the Evaluation Research Society Network (ERS) standards for program evaluation, and the collective expertise of The NETWORK staff in program evaluation.

### Who is the handbook for?

The handbook is designed for a variety of audiences including:

- Project directors who have decided to carry out some of their program evaluation tasks;
- People who will contract with an outside evaluator but who want to understand evaluation methods;
- Project staff who will work closely with project evaluators; and
- Grant writers who are writing an evaluation plan to submit with a proposal.

## What is its content?

Section 2 provides some ideas and theory that can help in thinking about evaluation. The reader will learn to determine the scope and direction of the evaluation process. Section 3 introduces the strategic steps to an evaluation and provides the reader with practical suggestions to prevent falling into evaluation traps. Section 4 helps the reader think about use of the evaluation report. Section 5 reviews some of the most common evaluation misconceptions. The handbook ends with a glossary of evaluation terms, references, and consumer tips on evaluation.

# SECTION 2

## ideas about evaluation: why? what? how?

### Why evaluate?

Before developing an evaluation plan you should step back for a moment to gain some perspective. What are your major purposes in evaluating your program? Are you most interested in improvement — in better understanding and running your program? Do you wish to make a case for your program's future, based on its effectiveness?

These questions lead to the two major categories of evaluation cited by most evaluation literature. Evaluators refer to them as *formative evaluation* and *summative evaluation*. Evaluation theorist Michael Scriven (1967) has developed this useful distinction between formative and summative evaluation. An instructional sequence that is evaluated to improve the sequence itself is known as *formative evaluation*. When the evaluation appraises the worth of the completed program, it is *summative evaluation*. As you carry out your project, both types of evaluation should be conducted.

While you may have many particular reasons for wanting to evaluate, your evaluation will occur in the context of (1) priorities for improvement of the program itself, (2) results for the community on which your program impacts, and (3) funding criteria set by the sponsors or the funding agency.

As you examine the ten reasons to evaluate a program listed below, you may find that your answer to the question — Why evaluate? — will include several of these reasons:

- to identify the program's most valuable activities ;
- to make cost- and service-effective management decisions ;
- to serve clients better ;
- to improve credibility with the funding agency ;
- to make better use of limited resources;
- to improve overall program performance;

- to give direction to the program;
- to increase future funding;
- to show success; and
- to raise staff morale.

## Evaluate what?

Now that you have decided the **why**'s, let's look at the **what**'s. Evaluation provides important learning about your program. Some of this information will appear as data collected in the different areas of your program. The data will either confirm your views regarding the effectiveness of your program or point out problem areas that need attention. Since every program has many different program elements and activities with varying degrees of importance, this guide cannot tell you which activities or elements should be evaluated. What it can provide are some considerations on how to select the activities you will evaluate. Because you do not have unlimited resources, it will help you to figure out which activities are worth evaluating more than others.

You may choose one or two activities to evaluate; you may choose several activities to evaluate in combination; or you may choose several activities to evaluate in sequence. Bank and Snidman (1981) suggest three strategies to evaluate program activities when you have some freedom of choice.

- *Funnel strategy:* you evaluate a comprehensive program area and then use the results of that evaluation to identify more specific components to evaluate.
- *Troubleshooting strategy:* you evaluate activities which are of most concern to you or are causing problems of time and money.
- *Piggyback strategy:* you evaluate those activities that are or will be repeated throughout the life of the program.

## Evaluate how?

Once you have weighed and considered your resources and decided **what** it is that you want to evaluate, you must give some attention to **how** you are going to do it. Here are some initial steps to get the ball rolling:

1. *Look for the basic information that already exists on your program* — e.g., past evaluations, proposals, evaluation procedures, and funding agency timelines. These sources will help to give you a clear idea of your purpose in evaluation.
2. *Discuss evaluation strategies with staff.* Choose basic strategies and identify in-house and outside resources. Determine the scope of the evaluation.
3. *Identify external resources needed for the evaluation of your program.* Assess costs for professional evaluators. Collect references before securing a consultant..
4. *Meet with your outside evaluator and/or your staff and plan your evaluation.* Once you have selected an outside evaluator, develop a detailed schedule in the form of an activity chart, checklist, or matrix, outline evaluation stages, tasks/responsibilities for activities, and a time frame for their completion. Assign responsibilities for these tasks.
5. *Formulate a plan for your evaluation.* After establishing communications procedures with your evaluator and other key staff, proceed step by step with your evaluation plan.

# SECTION 3

## ten basic steps to evaluate your program

The ten steps outlined below constitute an evaluation plan. The point of the evaluation plan is to gather evaluation methods and data together in a chart that graphically shows how steps proceed and interrelate. The ten steps provide you with a structure in which to begin evaluation.

In 1985, The NETWORK, Inc. received a grant from the U.S. Dept. of Education to conduct a project called *Educational Equity for Northeastern Hispanic Adolescents: Research and Action.* Upon receiving the good news, the project director called a staff meeting to discuss the steps to be taken. The project evaluator, who was also present, was familiar with the proposed activities for the project because she had also written the evaluation plan submitted with the proposal. Since the project director believed that evaluation is an integral part of any project, she wanted evaluation activities to start as the project began.

After a great deal of discussion concerning action steps to get the project underway, the project evaluator detailed her evaluation strategies. "We will start with our goals and our objectives," she said.

(As the evaluation steps proceed, you will see how the project, *Educational Equity for Northeastern Hispanic Adolescents: Research and Action,* carried out its evaluation.)

The ten steps are:

1. Identify project goals.
2. State objectives.
3. Decide on the scope of your evaluation.
4. Formulate the evaluation questions.
5. Design the evaluation.
6. Choose the method of information collection.
7. List evaluation activities and establish a timeline.
8. Estimate the cost of the evaluation.
9. Collect, analyze, and interpret the data.
10. Communicate and act on the evaluation findings.

# STEP #1: Formulate Goals

This step requires that you go back to the development of your project and critically examine the originally stated goals for your project. Have they changed as the project was implemented? Do they need rewording? Do they seem too broad or too ambitious? We recommend that you meet with your staff and review the implementation and outcome goals of your project. During this stage you may discover that you have goals that are too ambitious to accomplish or that you have expressed your goals too vaguely. Or you may find that some goals are fuzzy and may be interpreted in several ways. Sometimes it will be necessary to rephrase or reword goals to better reflect what is really happening in your project.

*Educational Equity for Northeastern Hispanic Adolescents* established the following goal:

TO INCREASE HISPANIC ADOLESCENT FEMALES' AWARENESS OF NONTRADITIONAL CAREER CHOICES

Notice that this goal is a broad statement of intent. It does not spell out how this awareness will be increased or by how much. "Nontraditional" career choices are not defined nor are there any indicators of success.

As a goal, it is fine. However, if you were to give the same goal to 20 different project directors, they would disagree on how to accomplish it. The next step — the how to do it — takes us to the objectives of a program.

Next, select the goals of the project that the evaluation will assess. Then, list these project goals according to priority. This procedure is based on the assumption that an evaluator from outside is coming into a program whose processes are essentially unknown. While official statements of the intentions of the program may exist, these may not be accurate models for what is really going on.

Listed below are some examples of stated goals for educational projects.

- To increase Hispanic adolescent females' awareness of nontraditional career choices.
- To increase the effectiveness of the math and science instruction delivered to students in the school district A.

Please note that the goals stated above are specific, practical goals. Presented with each one of these goals, you would be able to develop a set of objectives and a number of activities to carry them out. When goals are fuzzy, too broad, or too ambitious, the task of writing objectives and designing activities to carry out these objectives becomes overwhelming. For example, examine the following project goals encountered by this evaluator:

- To help parents help their children to learn.
- To reduce teenage pregnancy.
- To decrease drug abuse among adolescents.

There is a marked difference between these two sets of goals. The first set leads you to work; the second leaves you baffled. The goals in the second set are too broad, too vague, too overwhelming. Often an evaluator will start work by examining the goals of your project and may request that you clarify your goals before an evaluation design can be attempted. If the stated goals of your project are satisfactory, you should discuss with your evaluator which goals are your priority for the evaluation.

# STEP #2: State Objectives

Formulating clear and meaningful goals is the foundation of your evaluation design. Even more important is the statement of precise, clear, and measurable objectives. Your objectives are the destination of your program, the specific and verifiable aims. If they are precise and measurable, they provide the fundamental building blocks of your evaluation. Evaluating your

*Educational Equity For Northeastern Hispanic Adolescents* selected several objectives to carry out its goal and give the project director a sense of direction. Among these was:

AFTER SIX MONTHS OF USE OF THE MATERIALS BY TRAINED TEACHERS, THE EXPERIMENTAL GROUP WILL SHOW A SIGNIFICANT INCREASE OVER THE CONTROL GROUP IN THE LEVEL OF AWARENESS AND ACCEPTANCE OF NON-TRADITIONAL CAREER CHOICES AS MEASURED BY A CAREER INVENTORY INSTRUMENT TO BE ADMINISTERED AS A PRE/POST TEST.

While we decided to produce a video and an accompanying manual, other strategies to achieve the results we wanted (as expressed by our objective) might also have been appropriate. Next, we listed under the objective the various activities needed to carry it out. We selected the following activities:

Activity #1: To search out a movie director/producer experienced in issues related to women and minorities who would make the film within the allocated budget. This involved referrals, interviews, examination of portfolios.

Activity #2: To survey the population in order to determine the video content. This involved connecting schools that fit our criteria (concentration of Hispanics, geographic location, commitment of administrators and staff to work with us), developing a survey instrument, administering the survey, analyzing and interpreting results.

Activity #3: To write a script for the movie and develop the manual guided and informed by the findings of the survey questionnaire.

Activity #4: To develop the Career Inventory instrument. This involved examining instruments already developed in order to assess their appropriateness to our project. Because a good fit was not found, we decided to develop our own instrument. Instrument construction is a technical task that is expensive and time consuming.

Activity #5: To administer the Career Inventory instrument. This involved scheduling test administration at the six sites, duplicating the instrument, selecting which of the two schools in each city would be experimental or control, training two staff to assist the evaluator in instrument administration, traveling to the sites.

Activity #6: To train teachers in the use of the video and the manual. Once the video and manual were finalized, we started the training of teachers in the three experimental schools. Training teachers involved scheduling the training sessions, making arrangements for substitutes, establishing the desired performance, assessing the effectiveness of training.

Activity #7: To assist teachers in the implementation of the program. This involved traveling to the sites, observing classes, troubleshooting, implementation follow-up for six months.

Activity #8: To administer the posttest to control and experimental groups. This involved scheduling test administration and traveling to the six sites.

Activity #9: To analyze data and interpret the results. This involved comparison of pre/posttest for experimental and control groups using statistical procedures. A final evaluation report summarized results.

program can be seen as the documentation of how well you meet each of your objectives.

Now let's turn to what makes an objective worth its name. For example, "to lose weight" is not an objective but a general statement of intent, a goal. Objectives are statements of what will be different, by when, and how this difference will be measured. Objectives specify results, not strategies for getting results. Objectives are often confused with activities or strategies for getting a certain result. Using the same example, we can say that:

**To lose weight** is a goal.

**To lose 20 pounds by December 31** is an objective; and

**To take an aerobics class three times a week and to reduce calory intake to 1,000 calories a day** are activities or strategies you will use to get the result you want. The results of your program should be stated in your objectives.

Please examine the following "objectives" and try to answer the question which is at the heart of the objectives issue: "What will be different, by when, and how will this difference be measured?"

- To conduct 15 training sessions involving parents and students focusing on the effects of drug abuse.
- To develop a manual that acquaints parents with school procedures.
- To train teachers and aides to staff and manage writers' workshops.

These are no doubt worthy activities but they are *not* objectives. In order to be called objectives, they have to state the changes that occur as a result of the activity and establish indicators of success. How many individuals participated? What were the observable changes in knowledge, attitude, or behavior? What degree of change is acceptable as an indication that the expenditure of resources on the training or on the writing of the manual is justified? As you examine the evidence, ask yourself the question, *"What evidence would convince me that individuals are attaining the objectives we have set for this program?* Finally, a time frame must be established.

Writing objectives that are sound and at the same time realistic is not an easy job. Often, program people become irritated with the evaluator's insistence on measurable objectives and reply that "it is impossible to measure the effects of our program — we are dealing with things that can't be quantified." Measuring effects of a program, however, can be done in either a quantitative or a qualitative manner — often both are used together in the evaluation of a program. Once the distinction between objectives and activities is understood, writing objectives becomes easier. When writing objectives you need to avoid the following:

- **Objectives that are too vague and open to all kinds of interpretation**
  "To help parents help their children"
- **Objectives not worthy of their name**
  "To train 200 teachers in cooperative learning techniques"
  An obvious case of mistaken identity. Even if you trained 800 teachers, you would still be talking about an activity, a strategy to achieve a certain result. To be an objective, this would have to be rephrased to include (a) the result of the training that could be observed in the trainees; (b) the percentage of trainees who would show successful acquisition of the cooperative learning techniques; and, (c) the manner in which this change in knowledge/attitude is to be assessed.

Objectives are often confused with goals. "To increase awareness of nontraditional career choices," may be a clear and meaningful goal but it is *not* an objective. "To influence 80% of all seniors in the Bora-Bora School System to choose a nontraditional field of study upon graduation" is an objective. Why is it an appropriate objective?

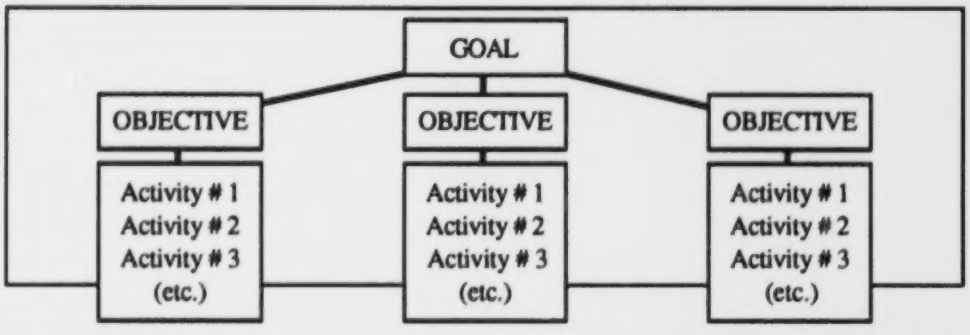
1. It names the target population — seniors in the Bora-Bora School District.
2. It includes measurement indicators — percentage of target population, number of students reached by program.
3. It indicates performance standards — how much of the measurement indicator will constitute success in meeting the objective.

Mark Ziegler (Contra Costa County, CA) suggests a mnemonic device for remembering the characteristics of clear objectives:

S pecific
M easurable
A ppropriate
R ealistic
T imebound

Once the goals, objectives, and activities are clearly stated, you are ready to work with your evaluator on an evaluation plan for your project. An evaluation plan states the goals, objectives, and activities that will achieve the results you desire. It states the indicators of successful accomplishment of each objective. It specifies the data collection procedures, the instruments to be utilized to gather data, the sources of data, and the methods and procedures to be used for the analysis of the data.

The graph below depicts the relationship among goals, objectives, and activities.




Listed on page 14 are questions to ask when you begin to formulate objectives, and on the following pages you will find a worksheet for the development of objectives. You may want to use those as an organizer for the basic information categories of your objectives.

## Basic Questions to Ask Yourself When Stating Program Objectives

1. What precisely do you intend to change through your program?
2. What would the funding agency accept as proof of your success?
3. What are your target populations?
4. What would happen to your target population without the benefit of your program?
5. What are the short term and long term effects of your program?
6. What side effects could result from your program?
7. Who might be affected by your program besides the target population?

## WRITING OBJECTIVES: WORKSHEET

Program Goal: To increase Hispanic adolescent female awareness of non-traditional career choices.

---

Target Population(s): Hispanic teenagers in grades 8-12 in schools in Boston MA, Hartford CT and New York City.

---

Objectives: • To survey a sample of the population in order to determine the substance of the video; • To develop a 30 minute video featuring six Hispanics in non-traditional occupations.

---

Activities:

Time Frame

#1: To search out a movie director/producer — April 30

#2: To develop a survey questionnaire — May 15

#3: To write a script for the video — July 31

#4: To review and revise the first draft — Sept. 30

#5: To schedule a showing of the video to a representative sample (150) — Oct. 30

#6: To pre/post assess the audience — Nov. 30

#7: To write a final report — Dec 31

---

Indicators of Success: • Internal reviewers' reaction to video (project staff, others at The NETWORK;

- External reviewers' reaction (representative sample);
- Audience's reaction as measured by a pre/post survey conducted at the time of the showing.

# WRITING OBJECTIVES: WORKSHEET

Program Goal:

---

Target Population(s):

---

Objectives:

---

Activities: Time Frame

---

Indicators of Success:

# STEP #3: Decide on the Scope of the Evaluation

After you have scrutinized and prioritized your goals and objectives, you need to decide how comprehensive and thorough the evaluation is going to be. The scope of the evaluation is determined by various factors. You will want to pay attention to at least three of these factors: *Resources, Size, and Skills.*

**Resources** refers to the amount of time and money needed to carry out the evaluation tasks. Here "fatal flaws" often arise. In their budgets, some project directors allocate only token funds to evaluation, mistakenly assuming that evaluation is cheap and quick. When they realize that a meaningful evaluation is neither, it may be too late to find additional funds. The project director will have to settle for a token evaluation — quick and cheap — and thus perpetuate the misconception that "evaluation does nothing for our program."

Time and money can quickly become entangled when the evaluator is not called upon until the program is almost over. At that point little can be done. The data needed to determine program effectiveness or establish meaningful comparisons may not have been collected, and the lack of an evaluation plan from the beginning has resulted in helter-skelter evaluation. Faced with this situation, the evaluator is severely impaired in attempts to show program impact. Once again, the project director may end up with a token evaluation, one that may even misrepresent the program.

For our project we decided on the scope of the evaluation as we wrote the proposal. Two factors were considered: First, the fact that the project's goal was to develop a product—a video featuring Hispanics in non traditional occupations—and, second, that the program's length—12 months—would not allow us to collect much information on the video's impact on the target population. Thus, our evaluation was, for the most part, formative so that project staff could be assured of the quality of the product. We decided that impact would be assessed at one showing of the video to approximately 150 people.

The decisions we made that determined the scope of the evaluation were:

1. We would survey approximately 500 Hispanic adolescents in three school districts (Boston, Hartford, and New York City) to gather the data to inform the development of the video.

2. We would assess the impact of the video by arranging one showing to an audience of approximately 150 people.

**Size:** The size of an evaluation is determined by the number of evaluation questions asked, the number and complexity of instruments used, the number of respondents, and the number of occasions for data collection. Often a trade-off must occur between how lean your evaluation can become (before it becomes trivial) and the amount of resources available. For example, if your program serves 300 people but you lack the resources to interview all three hundred, you may decide to interview half, or a smaller sample, of your 300 participants.

The trick in sampling the population is to obtain a small group of people who are a microcosm of the larger population so that you can generalize the data obtained from a smaller sample to the larger population. There is no magic number that is considered adequate. If your population is small — 50 or under — it is advisable to use the entire population. For a

population up to a thousand, experts consider that 30% of the population is adequate. Above a thousand, the percentage needed for a sample decreases markedly (Fowler, 1984).

Besides representativeness of the sample, the other "fatal flaw" you want to avoid is sampling bias. If when administering a survey, you ask for volunteers, for example, you end up with only those who choose to respond. If you mail out survey questionnaires and only consider the answers of those who sent the questionnaires back to you, again, you may find that only those who felt strongly enough about the program answered. In either case, your sample is biased, and therefore, not representative of the whole population.

Methods for selecting a sample, regardless of the size of the population, are summarized below:

1. *Random Sampling*. Each member of a population has the same chance of being included. First, identify your population. For example, all Hispanic teenage girls in grades 11 and 12 in the school district. Then, draw names at random.
2. *Stratified Random Sampling*. This is useful when it appears that particular factors within the population might influence an individual's reaction to the evaluation measures. Examples of such factors are age, gender, occupation, geographic location. To obtain a stratified random sample, classify your population into categories and then randomly select individuals from each category.
3. *Matrix Sampling*. Not only are individuals randomly selected, but the items in the instruments are assigned randomly, so that each individual answers only a few items. For this sampling method, you need a larger sample and cannot make comparisons among individuals.
4. *Convenience Sampling*. In this method you interview all those individuals who are at a certain place, for example, all those you meet in the school cafeteria on a given day or time. This is a useful method for informal evaluations but you must be cautious when trying to interpret your data or make generalizations.

We made the decision to adopt a stratified random sampling procedure for the administration of the survey for the following reasons:

1. To be cost-effective and reduce travel, our strategy was to select schools with a heavy concentration of Hispanic students;

2. We were interested in adolescents 12 and older;

3. We wanted to be able to compare survey results along different lines—Hispanics in NYC, an "older" population, and recent arrivals in Boston or Hartford;

4. We included a vocational school because we were interested in a comparison of the views of high school adolescents with those held by vocational school students;

5. We selected a non-Hispanic control group in order to compare the views of different groups regarding career choices and occupations.

Graphically, our sampling plan looked like this:

| SEX | AGE | GRADE | TYPE OF SCHOOL | GEOGRAPHIC LOCATION |
|---|---|---|---|---|
| 50% male | 13, 14, 15 | 8 - 12 | high school | Boston – two schools |
| 50% female | 16, 17, 18 | | vocational school | Hartford – two schools |
| | | | | NYC – two schools, one vocational |

5. *Judgment Sampling*. Your sample is those individuals you believe can give you good information. Or you may select those individuals who you believe are typical of a type or an opinion existing within a larger population. Naturally, the interpretation from convenience and judgment sampling must be guarded because you do not know if the group really represents your population.

Skills: Three types of skills are required to conduct an evaluation. They are:

1. *conceptual:* select the evaluation questions, select the design that will provide answers to those questions, anticipate alternatives.
2. *technical:* construct or choose appropriate instruments for data collection, determine occasions and procedures for data collection, analyze and interpret the data.
3. *managerial:* assign responsibilities, make timelines, create budgets, keep individuals on track.

Concerning the scope of the evaluation, another decision involves the number of evaluation questions, the number of instruments, and how often these instruments are administered. A modest evaluation may include one or two questions and one instrument administered to a large number of respondents on one occasion. A moderate evaluation might ask three or four questions of a small number of respondents. An ambitious evaluation might ask many evaluation questions, using many instruments and administered on several occasions to a large number of respondents. Whatever you decide, remember that an ambitious evaluation cannot be conducted with modest funds. If your funds are modest, you are better off conducting a modest — but sound — evaluation, rather than attempting to do too much. Here, less is definitely more.

You are the person who must make the decision concerning the scope of your evaluation. You are aware of the constraints of budgetary considerations and the requirements of the funding agency. You know what resources you can use, the skills you have in-house and those that must be found outside. Whatever the scope of your evaluation, a good job is a good job, and you should settle for no less.

# STEP #4: State the Evaluation Questions

Central to an evaluation are the questions that will be answered. Your evaluation should be organized so as to answer the important questions about your program. Your choice of questions is crucial in the development of your evaluation plan.

Once you have selected the activity or activities to evaluate, you must work with your evaluator to choose a set of questions that will guide your evaluation. Below are some guidelines for the development of evaluation questions:

Objectives-based orientation, first proposed by Ralph Tyler, is concerned with the gap between program objectives and program outcomes. Using the author's program as an example, evaluators following this model would ask questions such as:

1. What percentage of Puerto Rican teenage girls changed their perceptions about equitable practices in the classroom?
2. What behaviors were observed that showed this change of attitudes?

**Responsive orientation**, associated with Robert Stake and many British evaluators, is concerned with responding to the issues and concerns of the individuals affected by a program (Stoke, *et al.*, 1975). An evaluator following this orientation would ask questions such as:

1. How do inequitable practices related to gender affect the day-to-day life of the target population?
2. What is the most crucial area in which these unfair practices must change and what can the program do to speed this change?

**Decision-making orientation**, associated with Daniel Stufflebeam (1971), regards evaluation as a precursor to action. Evaluators following this orientation would state their evaluation questions only after finding out who makes decisions about the program, what decisions must be made, what information is needed to make these decisions, what types of information the decisionmaker needs and in what forms. An evaluator would ask questions such as:

1. What information do administrators need in order to move forward with a program intended to increase the opportunities for Hispanic females?
2. What indicators of success would administrators need in order to extend the program to other sites in the district?

These are only some of the questions that are suggested by the program. Even though the approaches are different, the questions are equally effective. Most evaluators will use a more eclectic approach and try to address all these areas. Again, the scope of the evaluation is the key factor regarding how many questions you will ask, the areas you will address, and the comprehensiveness of the answers that you will receive.

Many programs fall into three categories: product development, service provision, and training programs. Here are some examples of evaluation questions appropriate to each of these types of project:

Product Development Project: Sample Evaluation Questions (Used for the evaluation of the *Educational Equity for Northeastern Hispanic Adolescents* project)

1. Was the product developed in a timely fashion?
2. What steps were taken to ensure the quality of the product?
3. What information was collected to guide and inform the development of the product?
4. How many people will be exposed/have access to the product? Who are the potential users?
5. How will the product be disseminated?
6. What information is available/was collected to determine the impact of the product on the target population?

Service Provision Project: Sample Evaluation Questions

1. How were the services selected?
2. What evidence is there that the services provided are needed and relevant?
3. How many people benefited from the services the project provided?
4. How did the recipients rate the services provided?

5. What information was collected/is available to determine that the services provided filled a gap or ameliorated existing problems and situations?

Training Program Project: Sample Evaluation Questions

1. What needs prompted the training?
2. How was the format and content of the training proposal/sessions selected?
3. How many people were trained?
4. What assessment was made of participants' pre-training condition?
5. How was the impact of training on participants' reactions, attitudes, knowledge, and behaviors assessed?

# STEP #5: Design the Evaluation

Now that you have specified your goals and objectives, and realistically assessed the situation so that the scope of your evaluation is manageable but meaningful, you are prepared to design your evaluation.

The term "design" applies to evaluation in two different ways. We can speak of "designing" an evaluation in the sense of putting all the component parts together into a coherent plan. We can also speak of "evaluation designs." Designs are structured ways of assembling the variables of a program logically and graphically to measure the program's success. These designs often use the technique of comparison in one form or another. True experimental controls, however, are rarely possible except in certain research projects. Finding a way to isolate the effects of your program from everything else that has impact on your target population and then to measure that effect is difficult.

As you have already seen in examining the problems of goals and objectives (Steps # 1 & 2), an evaluation design is not as simple as the ten basic steps may have implied. The variables that comprise the program — people, activities, strategies, and materials — will not all be causes of the program's significant effects. The evaluator must isolate those program features that are essential and characteristic. A good deal of what goes on in the program may be irrelevant to its impact.

The evaluator must measure not only how much a program has succeeded but in what ways. What went on that worked? How could other programs adopt these strategies to achieve similar results? What program components would function well in other settings? What spinoffs or unanticipated outcomes were observed?

All these issues add complexity to the task of designing an evaluation. While the ten steps could serve as a backbone for an approach to evaluation, they are not the ultimate answer. You and your evaluator must seek those answers. In the definitive reference work on the subject by Donald T. Campbell and Julian C. Stanley (1971), the authors list 12 "Factors Jeopardizing Internal and External Validity" of designs. These factors are the interfering effects and imperfections found in real life. Below are the most common of these factors:

- **History:** "Outside world" events may influence the success of a program.
- **Maturation:** The participants in a program are undergoing personal maturation apart from what they are learning in the program.

- **Testing:** The test itself may influence participant's responses to a subsequent test.
- **Instrumentation:** Change in instruments, tests, or procedures can result in changes in results which may be seen as changes produced by the program.
- **Selection:** Working only with volunteers or allowing people to self-select into the program would add the factors of motivation and interest to the effect of the program.

After doing her "homework" (steps 1 - 4), the project evaluator was ready to present the project director with a detailed evaluation design. The evaluation design looked like this:

| ACTIVITYY | EVALUATION QUESTIONS | SOURCE OF INFORMATION | METHOD/ PROCEDURE | STAFF RESPONSIBLE | DUE DATE |
|---|---|---|---|---|---|
| Development of video | • What information must we collect to ensure that the video is relevant to the target population? | • Hispanic adolescents<br>• Hispanic parents<br>• Counselors<br>• Members of the community | • Survey<br>• Interviews<br>• Review of literature | • Project evaluator<br>• Project staff | December 1987 |
| | • What experience and characteristics does the producer need in order to produce the video? | • Letters of reference<br>• Examination of relevant prior work | • Interviews | • Project director | January 1988 |
| Selection of role models | • How will the role models be selected? | • Referrals | • Contact agencies<br>• Telephone | • Project director | January 1988 |
| Video script | • How will the script be developed? | • Survey results | • Conferences with project staff and with movie producer | • Project director | January 1988 |
| Impact of video | • How will the audience be selected to ensure representativeness? | • Referrals<br>• Master list | • Stratified sampling | • Project director<br>• Project evaluator | March 1988 |
| | • How will the audience reaction to the video be documented?<br>• What changes can be detected in knowledge and opinion? | • Members of the audience | • Pre/post test | • Project director<br>• Project evaluator | May 1988 |

The charts that follow are summaries of evaluation designs for a research project, a product development project, and a service project.

BEST COPY AVAILABLE

# EVALUATION MODEL FOR A RESEARCH PROJECT

## Model:

Evaluation Research (Campbell 1969; Cooley 1975)

## Project:

**The Bilingual Validation Study**

The NETWORK, Inc., conducted a study that focused on the quality of existing bilingual/ESL programs. It was hoped that, as a result of close scrutiny of a representative sample of existing bilingual/ESL programs, exemplary practices would emerge and serve as a model for programs across the country.

The evaluation focused on project monitoring and on answering to the two evaluation questions below.

## Evaluation:

- Focuses on explanation of educational effects

## Evaluation Questions:

- What is the contribution that this study will bring to the field?
- Is the study methodologically sound?

## Action steps:

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| Assess pre-conditions<br><br>State the need (rationale) for the study<br><br>Clarify the benefits that the findings will bring to the field | Plan the study<br><br>Record the plan and its monitoring system | Prepare or select instruments<br><br>Decide on what data to collect<br><br>Decide on when to collect the data | Select sample<br><br>Decide on sampling plan<br><br>Form control and experimental groups | Decide on methods of analysis | Collect data<br><br>Administer instruments | Analyze the results<br><br>Interpret the results | Write final report<br><br>Describe program's goals and objectives; data collection methods and timelines; instrument selection/ development; include copies of instruments; data analysis procedures; report findings and make recommendations.<br><br>Disseminate findings. |

Campbell, D. "Reforms As Experiments," *American Psychologist*, 1969, 409-429.
Cooley, W.W., & P.R.Lohnes. *Evaluation Research in Education*. New York: Irvingston Publishers, 1975.

# EVALUATION MODEL FOR A PRODUCT DEVELOPMENT PROJECT

## Model:

Transactional Evaluation (Stake 1975; Rippey 1971)

## Project:

**Reducing Barriers for Hispanics**

The purpose of the project was to produce and test a package of materials in Spanish and English for counselors, students, and parents that would promote sex equity for Caribbean Hispanic women and girls, by identifying and reducing internal and external barriers to sex equity.

## Evaluation:

- Describes process for the development of the project

## Evaluation Questions:

- Is the description of the process honest, straightforward, and detailed enough to enable someone else to develop a similar product?
- What product revisions are needed to make it more effective in reaching the target population?

## Action steps:

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Describe the usefulness of the product to be developed | Develop first draft of product | Develop second draft | Test product with sample of target population | Write final report | Identify target audiences |
| Describe characteristics of target population | In-house review | In-house review | Revise | Describe the process (design, development, revision, field testing) | Disseminate product |
| Describe need for product | Revise | Revise | Finalize product | Report results of field testing | |
| | Finalize first draft | Outside review | | | |
| | | Revise | | | |
| | | (Repeat Step 3 or parts of it as needed) | | | |

Rippey, R.M., (Ed) *Studies in Transactional Evaluation*. Berkeley, CA: McCutchan, 1971.
Stake, R.E., et al. *Evaluating the Arts in Education: A Responsive Approach*. Columbus, OH: Charles E. Merrill, 1975.

# EVALUATION MODEL FOR A SERVICE PROJECT

## Model:

Goal-oriented Evaluation (Bloom 1971; Provus 1971)

## Project:

**Leadership Training for New York City School Restructuring**

The purpose of this project is to train facilitators and opinion leaders in school-based management and shared decision making. The project will: (a) assist schools that want to begin exploring school-based management; (b) conduct facilitator and opinion leader orientation and training retreats and orientation sessions; and, (c) provide professional development activities for school teams in the design, planning, and implementation of school-based management models.

## Evaluation:

- Assesses program effects
- Assesses effectiveness of services provided
- Describes changes that can be attributed to program activities
- Assesses program's success in achieving desired outcomes

## Evaluation Questions:

- What do participants *know* now that they did not know before their involvement in the program?
- What *attitudes* has the program modified?
- What do participants *do* now that they could not do before their involvement in the program?

## Action steps:

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| Critically examine program's major goals<br><br>Conduct needs assessment | State program objectives in terms of desired outcomes | Find out as much as you can about the program<br><br>Activities: frequency, duration, special features<br><br>Target population: number, age, characteristics, needs, attrition<br><br>Staff: characteristics, roles, responsibilities | Collect data to determine whether objectives were attained<br><br>Develop data collection plan: *when*, *how* data will be collected, by *whom*, and *who* the respondents are<br><br>Develop/select instruments<br><br>Administer instruments<br><br>Score instruments | Analyze data | Write final report<br><br>Describe program's goals and objectives; data collection methods and timelines; instrument selection/development; include copies of all instruments; data analysis procedures; report findings and make recommendations |

Bloom, B.S., J.T. Hastings, and G.F. Madaus. *Handbook on Formative and Summative Evaluation of Student Learning* New York: McGraw-Hill, 1971.

Provus, M. *Discrepancy Evaluation for Educational Program Improvement & Assessment*. Berkeley: McCutchan, 1971.

# STEP #6: Choose the Methods of Information Collection

Various methods exist to collect information for evaluating your program. They include questionnaires, interviews, observations, tests of knowledge or skills, and existing records. The advantages and disadvantages of each method, along with some planning guidelines, will be discussed in this section.

How do you decide which data collection methods to use? The choice depends upon the specific evaluation questions, the time available to collect the information, and the amount of money and staff available to do the data collection. Often, too, it will be necessary to use more than one type of collection method. The general guidelines for each type of data collection device will enable you to begin planning how to collect information. (See the chart on the following page.)

The first question concerning instruments is whether to use one that has been developed by others or to construct your own. No clear-cut answer to this question exists. Instruments developed by others may have been developed with different purposes and different types of respondents in mind. What works for them may not work for you. On the other hand, constructing your own instruments may be a difficult task, especially if you lack the expertise to do it. Some instruments can be modified to fit your needs, certainly an easier task than trying to start from scratch. The thing to remember is that the purpose of an instrument is to collect information that will answer your question or questions. While a good instrument does that, a poorly constructed one does not. Textbooks and references for instrument construction as well as consultants are available. Use them as needed.

The purpose of this section is not to enable you to develop instruments to evaluate the effects of your program. Rather, it should serve as an organizer to help you think more clearly about the task. Before you make a decision about whether to use an existing instrument, develop your own, or hire a consultant, here are a few points to consider:

1. *What do I want to learn from this instrument?* Clarifying the purpose of the instrument is more difficult than it sounds. In order to develop a good instrument, you must know exactly what it is for.
2. *How long does it take to construct an instrument?* Longer than you think. After you conceptualize the instrument, you need to write the items. The items have to be reviewed by someone familiar with the subject and the respondents. A second draft follows a first draft, and often a third draft is needed.
3. *How good does the instrument have to be?* An instrument that will be used to collect information about your program must be of high quality. It cannot be poorly constructed and it should not be trivial (e.g., the famous "happiness sheet"). To assure that it is valid, reliable, and provides meaningful information, you will need expert assistance.
4. *How will I use the information gathered?* The time to decide how you will use information is while you are developing the instrument, not after. If you wait until you have developed the instrument to ask yourself that question, you may end up with useless data.

The chart below is adapted from Bank and Snidman's *Guidebook for Evaluating Dissemination Activities* (1981). It summarizes the following discussion on types of instruments and provides some help in comparing instruments.

## Comparison of Instruments and Procedures

| | Advantages | Disadvantages | Use in Evaluating |
|---|---|---|---|
| Tests of Knowledge, Skill | • Are direct measures<br>• Can be used to get reliable comparative data<br>• Usually have high credibility<br>• Can be re-used | • High development costs necessitated by extensive pilot-testing and revision<br>• Some technical skills needed for development | • Awareness sessions<br>• Training sessions<br>• Program follow-up |
| Questionnaires, Interviews, Discussions | • Can reach large number of people quickly, efficiently<br>• Usually have moderate to high credibility<br>• Can provide new insights<br>• Can be re-used | • Moderate to high development cost necessitated by moderate amount of pilot-testing and revision<br>• Interviews and discussions require trained personnel<br>• Follow-up procedures for questionnaires may be costly<br>• Findings sometimes biased by irrelevant factors (respondents' ability to read, idiosyncratic events, etc.) | • All aspects of a program |
| Observations: Structured/ Impressionistic | • Can provide background and details for other data<br>• Usually have high credibility if done thoroughly<br>• Rich source of anecdotal data | • High development costs of structured observations necessitated by extensive pilot-testing<br>• Require trained observers<br>• Time-consuming for data gatherer<br>• Time consuming for data analysts | • Awareness sessions<br>• Training sessions<br>• Program implementation and activities |
| Documentation | • Usually easy to set up<br>• Usually easy to monitor and collect<br>• Rich source of details | • Time-consuming for respondents<br>• Potential high level of irrelevancy | • All aspects of a program |

The remainder of this section will briefly discuss what each type of instrument can and cannot do, and give you some tips for their construction.

## Questionnaires

The *questionnaire* presents each individual in the sample with some type of information and requires some manner of written response. For younger children or others who have difficulty reading, the question may be in the form of a picture and require the respondent to answer by selecting another picture or drawing. Questionnaire items may be open-ended (requiring the respondent to write in some type of answer) or closed (in which case the respondent selects an answer from choices provided). The purpose of the questionnaire is to collect certain types of factual information or attitudes of those responding.

**Advantages and Disadvantages of Questionnaires**

There are six major *advantages* to the questionnaire, three of which have to do with time. First, the questionnaire can be administered to the entire sample at once. Second, the questionnaire can be mailed, reducing the administration time. Third, the respondent can take as much time as necessary to think about the responses.

The fourth advantage to the questionnaire is that respondents can remain anonymous. Often individuals will respond more honestly if they do not have use their names. Fifth, since each individual is asked the identical question in the same format, the format of the questionnaire ensures a greater level of uniformity in responses than does an interview. Additionally, with the closed format questionnaire, uniformity within possible answers will also exist, making the tabulation and interpretation of responses much easier. This may not be true of the open-ended format because it results in a wide variety of answers.

The sixth advantage of the questionnaire is that the data are more easily summarized and interpreted than with the interview. This is especially true for a closed questionnaire. The analysis of open-ended questionnaires may be more difficult. But, if the questions are well-stated, this is still less of a problem than with the interview.

There are five major *disadvantages* to the questionnaire. First, little flexibility is possible. During an interview the interviewer can pursue interesting lines of thought or clarify unclear issues. On a questionnaire, if something on a questionnaire is unclear to the respondent, it cannot be clarified. Nor can interesting areas be investigated or unclear answers expanded upon.

Second, many people have difficulty in reading or expressing ideas in writing. If the questionnaire is not well-written, the respondent will have difficulty answering. Similarly, if an open-ended questionnaire is used, those who have difficulty expressing ideas in writing may not be able to respond accurately and completely. A third disadvantage of the questionnaire is the lack of rapport with the respondent. Developing a good rapport with respondents often helps draw out honest and complete responses.

The last two disadvantages of the questionnaire have to do with obtaining responses. Because questionnaires are sent through the mail, the number of returns may be low. Even after repeated follow-ups, many respondents may not return the questionnaires. Finally, those who do respond may be considered a biased group. That is, biased data may result if

those who do respond to a questionnaire are different from those who choose not to respond.

**Guidelines for Development**

A. When planning the questionnaire:
   1. Identify the topic of the questionnaire. Decide on this topic before beginning to develop the questions and stay within that area.
   2. Determine the intended audience before the questions are developed. Audience characteristics will affect the format of the entire questionnaire and the phrasing of each question.
   3. Ask yourself and other members of your staff what you want to find out from the respondents.
   4. Determine ahead of time how the questionnaire results will be analyzed, because this will affect the format of the questions. For example, if the responses are going to be machine scored, open-ended questions cannot easily be used. Or, the closed format might be used and respondents asked to answer on a separate sheet which could then be scored directly by the computer. The questionnaire should also be arranged to facilitate scoring responses by grouping similar items together (i.e., all yes/no type questions together).
   5. Ask only for information which cannot be obtained elsewhere. The purpose of the questionnaire is to collect some type of information or attitudes from each individual. Each person is responding because their input is necessary and of interest.
   6. Keep the questionnaire short. A person is much less likely to respond to a long questionnaire.

B. When developing the questions to include on the questionnaire:
   1. Order the questions in a logical manner. Start with the most general types of questions; then move on to the specifics.
   2. Limit each question to one idea. If a question does cover more than one point, it is impossible to interpret the results.
   3. Do not ask leading questions that lead the respondent toward the desired response.
   4. Word each question as simply and clearly as possible. Do not include information that is unnecessary to the question. Avoid technical terms, unless they are appropriate to the audience.
   5. Include a definite point of reference to ensure that each individual responds to the same questions. For example, if a question asked, "How many hours do you work?" respondents may answer in terms of hours per day, per week, or in other ways. Changing the question to ask "How many hours per day do you work?" ensures similar responses.
   6. If a closed-question format is used, try to include options that cover all possible aspects of that question. Do not limit the answers to only one side or part of an issue. Additionally, since it is often difficult to anticipate all possible choices to include in a closed format, use the category "other" and allow a space for the person to write in a

response.

C. When planning for administration of the questionnaire:
   1. Include complete, clear directions on how to respond to the questionnaire. Explain exactly how to respond and where. Do not leave anything to the respondent's imagination.
   2. Include a cover letter with the questionnaire. A personally addressed letter will help to establish rapport with the respondent, to explain why the questionnaire is being sent, and to encourage its return. Be sure to state a deadline for return.
   3. Ensure that the questionnaire and all other correspondence is neat and easy to read.
   4. Include a stamped, self-addressed envelope to encourage returns.
   5. Use postcards or other means to follow up on non-responses. Remember that the more questionnaires returned, the less biased the sample.

## Interviews

You can conduct interviews either by phone or face to face. You can interview individuals or groups. The interviewer obtains information by asking questions and posing problems and then listens to or records the answers.

One of the questionnaires developed by the project evaluator was designed to assess Hispanic teenagers' perceptions of themselves regarding academic matters. Some of the items are listed below.

***Examples of Questionnaire Items:***

***DICHOTOMOUS QUESTIONS OR LISTS***

1. At my house, boys and girls get the same amount of encouragement and support concerning school matters. YES NO
2. My family would be proud of me if I got a job, even though I don't finish school. YES NO

***RANKINGS***

1. Number in order (5 being the highest) what your family believes to be more important for you:

_____ Getting a good job.
_____ Getting married.
_____ Finishing my high school education.

***Sample Items of Teacher Survey Questionnaire***

***RATINGS***

| | Strongly Agree | Agree | Disagree | Strongly Disagree |
|---|---|---|---|---|
| 1. Hispanic boys and girls face the same pressures (family, social) regarding educational opportunity as other students do. | 1 | 2 | 3 | 4 |
| 2. Hispanics as a group do not have the same commitment to work and education as other Americans do. | 1 | 2 | 3 | 4 |
| 3. The major barrier to success of Hispanics is the language. | 1 | 2 | 3 | 4 |

The *interview* is basically an oral administration of a questionnaire. Rather than requiring the respondent to read a question and answer in writing, the interviewer asks the question and takes notes on the respondent's oral response. The interviewer may choose to follow up an individual's response to obtain further information. Later the notes from the interview are summarized.

A structured interview makes use of an interview

BEST COPY AVAILABLE

protocol which must be prepared in advance. The protocol explains the purpose of the interview and asks respondents the same questions in the same order. The interviewer probes to clarify misunderstandings or ambiguities. This approach—as opposed to the unstructured interview format—minimizes interviewer bias and makes analysis of the answers easier.

**Advantages and Disadvantages of Interviews**

Seven *advantages* to the interview all relate to the human interaction element of the interview process. First, questions that may be difficult to phrase in writing are often easier to express and explain orally. Similarly, a second advantage of the interview is that it eliminates the dependence upon reading skills, making it ideal for non-readers, young children, non-English speakers, or other groups where the ability to read is limited.

A third advantage of the interview it that is allows for the establishment of rapport between those involved. This can, in turn, increase the honesty and depth of answers. A fourth advantage is that the interviewer has the opportunity to observe attitudes expressed through body language and voice inflection. Non-verbal communication sometimes gives clues about how strongly an individual actually feels about the question.

The fifth advantage of the interview is its flexibility. Unlike the questionnaire, if individuals do not understand some of the questions or terms being used, they can ask for further clarification. Questions or terms can be rephrased until the communication is clear. A sixth advantage of the interview is that responses can be pursued until the interviewer obtains enough information. While the questionnaire stops collecting information once the questions are answered, an interviewer can follow up on a respondent's answers until satisfied.

Finally, the return rate of the interview is not usually an issue, and there is seldom a need to track down missing parts of the sample.

There are five major *disadvantages* to the use of interviews, four of which relate to that same human interaction element. First, results of the interview rely strongly on the interviewer's interpersonal skills and communication ability. This advantage can just as easily turn out to be a disadvantage. The interviewer may threaten some people, unknowingly lead responses in a certain direction, or just fail to communicate effectively.

A second disadvantage of the interview is that each respondent may not be asked the same question. The interviewer, unless highly trained, may get off track, omit questions, or alter the meaning of a question by slightly rephrasing a few words. Third, since the interview requires the presence of the interviewer at every session, the interview is a more expensive and time-consuming way to collect information. This leads directly to the fourth disadvantage—having to use smaller sample sizes because of the increased cost and time factors.

The fifth disadvantage of the interview is that it is sometimes difficult to summarize and interpret. The analysis of the interview results must be carefully thought out ahead of time

and the actual summarization may require a significant amount of time.

**Interviews - Guidelines:**

A. When planning the interview:
   1. Identify the one topic of the interview.
   2. Determine the intended audience.
   3. Determine how the results will be analyzed.
   4. Ask only for information which cannot be obtained elsewhere.

B. When developing the questions to ask during the interview:
   1. Order the questions in a logical manner.
   2. Limit each question to one idea.
   3. Do not ask leading questions.
   4. Word each question as simply and as clearly as possible.
   5. Include a definite point of reference to ensure that individuals respond to the same questions.

C. When planning for administration of the interview:
   1. Establish a method for recording interviewee responses so that the actual wording is preserved as closely as possible. Recording methods include taking notes during or after the interview, using preplanned recording sheets, or tape-recording the session.
   2. Train the individuals conducting the interview. They should be able to conduct each interview in the same manner, ask the same questions in the same order, and avoid any emotional responses to the interviewee's answers which might affect future responses. The interviewer must be trained to probe for additional information and accurately note responses to each question. The interviewer should also be trained to use the first few minutes of the interview to establish a good rapport with the interviewee by explaining the purpose of the interview.

The following pages give an example of an interview protocol, developed for *Educational Equity for Northeastern Hispanic Adolescents: Research and Action.*

# Sample Interview Form

**FOR THE INTERVIEWER: GENERAL INSTRUCTIONS**

The interview format consists of the following sections:

A. Biographical data
B. Life story
C. Permission Agreement

The purpose of this interview is to record the life stories of adolescents who have been or are pregnant or parents. These stories will be used to create a book of stories for other adolescents about teen pregnancy and parenting. In order to elicit such stories, the interviewer and interviewee need to be comfortable and to feel comfortable with each other. To try to gain this atmosphere as quickly as possible, here are some suggestions:

- Introduce yourself.
- Provide some refreshments in the beginning.
- While interviewee is eating, chat about school, home, etc.
- When you are ready to begin, put away the food and discuss with the interviewee why you are here and what is the purpose of the interview. Emphasize the anonymity, the usefulness to others of their story. If you wish to read the introduction, that's okay; or you can also simply talk through the introduction.

We would like the interviewees to talk as much as possible. So we begin each section of the interview with the general question that we hope will solicit the type of information we are interested in. If the question results in an extended discussion, that's fine. Do not feel you have to cover all the topics or ask a lot of the questions if the interviewee is talking to the right topics.

In order to make the interview anonymous, each person's data sheet and tape will be numbered so that the tape will be identified only by the person's number.

Be sure to number the Biographical Data Form and the interview tape. Be sure to turn on the tape recorder and test that it is in working order.

# Sample Interview Form

**SECTION A: BIOGRAPHICAL DATA**

Name: _____ Number: _____

Ethnic/Racial Background: _____ Age: _____

Person is: 1. ☐ Pregnant ☐ Parent

2. ☐ Has other children

3. ☐ Is married ☐ Is single

4. Age at time of first pregnancy: _____

5. Age at time of second pregnancy: _____

6. ☐ Lives at home ☐ Does not live at home

7. ☐ Lives alone ☐ Lives with: _____

Address: _____

_____

Home telephone: _____

Program/school: _____

School contact person: _____ School grade: _____

Other schools attended: _____

_____

_____

Family: Mother: _____ Father: _____

Sibling(s) and their age(s): _____

_____

_____

_____

_____

## Sample Interview Form

**SECTION B: LIFE STORY**

Let's talk about your past first.

**[Open-ended question:]** When you found out you were pregnant, what happened and how did you feel?

**[Some other topics to cover:]**

Do you have sisters and brothers? Do they live at home?

What do your mother and father do? What about your sisters and brothers?

Do you go to church? What does it mean to you?

What future do you think your parents have wanted for you and your sisters and brothers?

Do you have friends or relatives who are also important to you? Who are they?

Did you have any plans for your future before your pregnancy?

When you discovered you were pregnant, who did you tell first?

What was your relationship with the father-to-be? What did the father say about the pregnancy?

Why did you decide to continue your pregnancy? Who did you discuss it with? What suggestions did they have?

Did you consider other options, such as adoption or abortion?

How did you learn about sex? In school? From friends? From parents?

Were you surprised when you became pregnant? Why? Did you use contraception? Why?

Did you know about contraception? If so, how did you learn?

Let's talk about your present life and situation.

**[Open-ended question:]** Describe a day in your life.

**[Other topics to cover:]**

Who do you live with? (Alone ___ ; Boyfriend ___ ; Father of child ___ ; Family ___ ; Other friends ___ ; Relatives ___ ) How is that working out?

What is (was) pregnancy like? Were you surprised? Was it different from what you expected?

Today, what are your future goals? How are you preparing to meet those goals? Who is helping you to attain those goals? (Friends ___ ; Parents ___ ; Church ___ ; Relatives ___ ; Child's father ___ )

How does or will parenting affect your goals?

Are you attending school? Do you like school? Do you plan to stay in school?

How did the school treat you during your pregnancy? If you are a parent, how does your school treat you?

What do your friends think about your pregnancy? Your family?

How are you getting along with your parents at this time?

Do you see the child's father? Does he participate in the child's birth? Does he participate in childrearing?

What are your childcare arrangements? Are they adequate?

Now let's talk about the future.

**[Open-ended question:]** Tell me what you would like for the future.

**[Other topics to cover:]**

What do you see as the future of your child (boy/girl)? Would that future be different if child was the other sex?

What do you see as your future? Do you want to continue your education? Do you think you will be able to reach the goals you stated before?

Would you like to marry? Now or later? Would you like to have more children? When?

How would you feel if your child became pregnant when she was a teenager? What actions would you recommend for a girl? What actions would you recommend for a boy?

As a final question, since you are an expert on teen pregnancy, what would you like to tell other teens that you think is important for them to know?

**[After the story section is over, thank the interviewee for their time and knowledge. Ask if they have any questions.]**

Now that we know some facts about you, would you please read and sign the permission agreement **[Read the release form, section C, aloud to the interviewee and ask if she/he has any questions. Be sure that they sign it.]**

If we decide to include some pictures of teens in the book, would you be willing to be pictured in it?

## Sample Interview Form

**SECTION C: PERMISSION AGREEMENT**

I hereby consent to the use of any personal data which has been obtained in a taped interview with the staff of The NETWORK, Inc. I understand that my interview will be used in the development of booklets and posters on the issues of teen pregnancy and parenting. These materials are being produced by The NETWORK, Inc., as part of a Women's Educational Equity Act project, funded by the United States Department of Education.

Confidentiality will be assured. Only pseudonyms will be used.

I agree to contact my parent/guardian for their permission.

I hereby release The NETWORK, Inc. and its licensees from any claim by me arising out of my statements from the interview or the use of them

Signed: _____

Date: _____

Name: _____ (print)

Address: _____

_____

_____

# Observations

*Observation* is a method of collecting information by watching what is going on at a certain time. The observer may be interested in whether certain behaviors occur or in existing characteristics of something. The observer may follow a set of predetermined guidelines when doing the observation, or may just take notes on what is happening.

Observation, as a technical term, is a procedure for gathering information without interfering with or affecting the situation being observed. Observation can occur just once or be conducted at regular intervals. It can be structured and make use of checklists, ratings, and scales, or be impressionistic.

If you are going to conduct structured observations, you should:

1. Prepare a checklist of the characteristics to be observed;
2. Train observers to distinguish when those characteristics are present; and
3. Select observation occasions to ensure representativeness.

If you are going to conduct impressionistic observations, remember that:

1. They may be influenced by the observer's mood, timing, or special events;
2. They require experienced and skilled observers;
3. They require some familiarity with what is being observed; and
4. They may give you a "feel" for the phenomenon being observed, but may not be a useful guide to further actions.

**Advantages and Disadvantages of Observations**

Four major *advantages* accrue from using observations to collect information. First, the observation is a useful method for collecting information that is difficult to obtain through other means. For example, it would be hard to use questionnaires or interviews to obtain accurate data on the number of times teachers used positive verbal reinforcement with their students, but an observer could count the actual occurrences.

A second advantage is that this method does not rely on recall of past events. So, using the observation as a technique to collect data is a more reliable method for certain types of behaviors or characteristics. The third advantage of using observations is that an individual's point of view does not affect the data collected. Questionnaires and interviews, by their nature, involve people's attitudes. This is appropriate where information about attitudes is of importance. However, where it

One of our projects was heavily dependent on whether teachers integrated the "critical thinking lesson" we had developed into their daily classroom activities. Since the impact of the program could not be assessed without implementation, the evaluation structured observations in the project schools, as follows:

Sit in the classroom for one class period, observe the teacher, and then place a checkmark if one of the listed behaviors occurs.

| | yes | sometimes | no |
|---|---|---|---|
| 1. Teacher integrates the "lesson" into his/her teaching e.g., makes it look like a natural part of the curriculum and not like an extra burden. | ____ | ____ | ____ |
| 2. Teacher communicates to students (verbal/non-verbal behaviors) that the "lesson" is important. | ____ | ____ | ____ |
| 3. Teacher encourages questions regarding the "lesson," explains, clarifies doubts. | ____ | ____ | ____ |

is more important to obtain objective information, observation is a good technique.

Fourth, observations can be used in a variety of settings and with many types of samples. For example, observations can be used to determine characteristics of facilities or to describe behaviors of individuals who cannot be interviewed or complete a questionnaire.

There are six *disadvantages* to the use of observations to collect information. First, the actual presence of an observer may alter the behavior of those being watched. Individuals consciously and unconsciously change their behavior when an outsider is present. Second, the observer may see behavior in a way that is different from those involved in the activity. An interaction in which a teacher is correcting a student may be seen as reinforcing and positive to the observer, while the student may perceive it as threatening.

A third disadvantage of the observation is the training required of the observer. While training observers is time-consuming and costly, reliable data can be obtained only with training. The fourth disadvantage lies in observer reliability. Behaviors can easily be overlooked or interpreted in different ways by different people. So, if there is more than one person acting as an observer, the reliability of the observations can be an issue.

Fifth, because the observer is required to be present at all times, observation is an expensive way to collect data. This leads to the final disadvantage of the observation—the need to use smaller sample sizes because of the time and cost factors.

**Observations - Guidelines:**

A. When planning the observation:
   1. Identify the category of behaviors on which data will be collected. The set of behaviors should be small enough to be reasonably observed during one session. Do not expect to collect information on every behavior at one time.
   2. Determine who will be observed. The sample will affect how the observations are done, the length of the observation, and the system for doing the actual observation.
   3. Decide ahead of time how the results of the observations will be analyzed. The data analysis can affect the format of the observer recording sheets and the types of information actually collected.
   4. Limit observation to areas in which information cannot be collected in other ways. For example, using observations to obtain information on students' ages would not be the best method of data collection. On the other hand, direct observation to determine equitable practices in the classroom would be appropriate.

B. When developing the observation instrument:
   1. Identify and clearly define each behavior that the observers will be looking for. The explanation of a behavior should not be open to interpretation by the observers. Having observers watch for disruptive behavior would result in very unreliable data because the term "disruptive" behavior is so vague. In comparison, observers could count the number of times a student left the seat.
   2. Develop a coding method, tally sheet, or other device that facilitates the observation process. If the observer has to take time out to write down words, the behaviors occurring during that time will be lost.

C. When planning to use the observation:
   1. Ensure that each observer is fully trained in the procedure. This would include an un-

derstanding of the definitions of each behavior, practice in using the device, and methods to be unobtrusive. If the purpose of the observation is simply to describe events as they occur, rather than watch for specific behaviors, the observer still needs to be trained in methods for recording behaviors.

2. When scheduling the observations, keep each period fairly short. Observing and recording behaviors is an intense activity that should be divided into several brief periods, rather than one long one. For example, if the observer needs to watch a classroom for a total of 30 minutes, ten three-minute series of observations would provide better data than three ten-minute periods. Of course, there may be times when the purpose of the observation is to describe what went on during an entire lesson, in which case it would not be possible to break up the observation periods.

## Existing Records

Use of *existing records* is a method of collecting information on the basis of what has already been compiled by others. This category would include any type of data which has been collected at a previous time or collected as a byproduct of other activities.

Existing records of documentation is a rich source of evaluation data. Documentations are not limited to records that are intentionally set up for a specific evaluation purpose. Some of the different types of records that you may want to collect and examine in order to document your program activities are:

| Narrative Records | Tabulated Records | Graphic Records |
|---|---|---|
| diaries, journals, and logs | travel forms | photographs |
| reports and memos | telephone logs | movies |
| anecdotes | attendance records | videotapes |
| minutes of meetings | expenditures | audiotapes |
| | sign-in sheets | artwork |

Documentation may be performed in a number of ways. It can be formal and comprehensive with clearly defined procedures, or it can be a selective review of documents to provide insights and leads for a more formal process. One thing to remember, in either case, is that the time to worry about documentation as an evaluation tool is NOT when the program is coming to an end. In order to document your program's activities, a recording system has to be in place from the beginning.

Using existing records is somewhat different from the previous three data collection methods. Existing records can be almost any type of information that has been compiled by others. Similarly, existing records can be byproducts of another activity, such as information on the number of times students leave the classroom for noninstructional activities. Other examples of existing records would include: student grades, attendance, test scores, completed work-books, sign-out sheets, letters, journals, account books, classroom tests, and final products done by students during a lab class. This list is obviously incomplete. Types of existing records are limited only by the fact that the data does need to be obtainable and fairly complete.

**Advantages and Disadvantages of Documentation**

There are five major *advantages* to the use of existing records. First, a wide variety of information is readily available. This category is so broad that some valuable data can be found for almost every situation. A second advantage is that they generally provide objective information. Of course, if student journals are used, the contents of each journal is subjective. But, in most cases, the records have been kept for some administrative purpose and have not been interpreted by others.

Third, existing records can be considered more credible than other types of information. Since they have been collected at the time of an event—for example, an attendance sheet—it is likely that the data are accurate. While a librarian may not be able to recall the number of books checked out during a certain day, records from the library check-out system can provide the information.

The fourth advantage of using existing records is the lower cost and shorter time required to obtain the information. This, in turn, leads to the fifth advantage, reduced staff time needed to collect the information. Unlike the interview or the observation, the use of records does not require extensive preparation time or the actual presence of a person during a long data collection activity. In fact, in many cases, the work can be done by someone not as highly trained.

There are four major *disadvantages* to the use of existing records. First, records may be incomplete. Some of the students may not have taken a certain test. Letters may have been lost or some account records misplaced. Since often no way exists to retrieve missing information, the resulting data may be incomplete.

Second, it may take some time and effort to extract the desired information from the existing records. For example, if journals are used, the necessary information would have to be compiled from many pages of information. Or, records that date back over a few years may have to be tracked down. The type of records used will obviously affect how easily obtainable the data is. The third disadvantage in the use of existing records involves legal requirements. In many cases permission must be obtained from each individual before the records can be seen or used in certain ways.

A final disadvantage can occur if people are asked to keep some type of records as a by-product of other activities. Unless the process is made very easy and the individuals approached very carefully, the recordkeeping may seem to be one more task to add to the workload. For example, one lab instructor may not mind keeping a tally of the number of times a piece of equipment is used, while another instructor may feel that it is an unfair burden because it requires the instructor's presence whenever the equipment is used.

**Existing Records - Guidelines**

A. When deciding to use existing records:
   1. Identify the area of information for the data collection; then select the appropriate types of existing records.
   2. Describe the sample from which the information will be collected as clearly and completely as possible (for example, Hispanic teenagers in junior high schools in three urban centers).

3. Decide exactly what type(s) of records will be used. Consider feasibility, cost, access, time, and legal implications in this determination.
4. Determine ahead of time how the information will be analyzed. The process used for data analysis will have implications for how the information is recorded.

B. When planning to collect information from existing records:

1. Identify exactly where the necessary information is located and whose permission must be obtained in order to gain access to these records.
2. Determine ahead of time any legal requirements which must be met in collecting or using the records.
3. Develop a method for extracting the necessary information and a means for recording that data. The method used must be easily understandable and consistent.
4. Train those who will be collecting the information. To obtain reliable data, each individual must record the same type of information. So, each person must clearly understand the categories on the recording forms and directions for their use.

Please examine the contact report and the workshop report that follow. They were designed to allow project staff to answer several questions at the end of the first year of a program: How many people called? Who were they? Where were they from? Who did they contact? What was the purpose of the contact? What next steps were taken?

**NECEA**

**CONTACT REPORT**
**SPRING 1990**

Staff: _____ _____

Date: _____

Phone: ❑ In Person: ❑

Gender: ❑ (70-2) N. Origin: ❑ (60-2)

Race: ❑ (50-2) Combined: ❑ (80-2)

Source of Referral: _____

**********

**SERVICES**

Planning: _____ Resource Request: _____

Follow-up: _____ Information: _____

Coordination: _____ Referral: _____

CLIENT: _____

TITLE: _____

SCHOOL BUILDING: _____

SCHOOL DISTRICT: _____

SEA/ORGANIZATION: _____

STATE: RI MA NH CT ME VT

REGIONAL: _____ NATIONAL: _____

ADDRESS: _____

_____

City: _____ Zip: _____

PHONE: ( ) - _____ EXT: _____

SUMMARY: _____

_____

_____

SUBSTANCE: _____

_____

_____

_____

_____

_____

_____

FOLLOW UP/NEXT STEPS: ❑ Info. Spec. ❑ AA's _____ ❑ Other: _____

Free Materials Sent: MM FT RF S.S. S.List CP others: _____

Information/Response Sent by Mail _____ Date

SEA Notified by Mail _____ Date

# NECEA WORKSHOP REPORT 1990

**LEA/Organization:** _____

**Site of Meeting:** _____

_____

**State:** _____

**Name of Presenter(s):** _____

_____

**Date of Event:** _____

**Staff/Consultant:** _____

_____

**Date of Report:** _____

**Gender** ❑ **(70-2) Race** ❑ **(50-2)**
**Combined** ❑ **(80-2)**
**National Origin** ❑ **(60-2)**

| Audience | Number |
|---|---|
| Teachers | |
| Administrators | |
| Parents | |
| Others | |
| Total: | |

**Title:** _____

_____

**Goals/Objectives:** _____

_____

_____

_____

_____

**Design Outline:** _____

_____

_____

_____

_____

==========

**Overall Evaluation:** _____

_____

_____

_____

_____

**Please attach the following and check:**

❑ **Handouts** ❑ **Evaluation Forms** ❑ **Attendance List** ❑ **Expense Report** ❑ **Materials Requests**

## Tests

The part of this section that deals with tests is very brief. This is not an oversight on our part. Rather, we have decided to keep it brief in part because a handbook such as this cannot adequately address the vast range of issues involved in test development. Developing tests is time consuming, expensive, and technical. If you find that you need to develop specific tests to assess the impact of your program activities—or if you need advice on selecting appropriate tests—talk to an evaluator. Test development is better left to experts.

Tests of knowledge acquisition can be written or oral, long or short, formal or informal. Tests can tell us about the respondents' recall of facts, events, or concepts. They can also inform us of the respondents' ability to apply knowledge to specific situations.

Tests of knowledge acquisition can include true/false items, multiple choice items, matching items, or essay questions. Often a test will contain more than one of these formats. You can use tests at awareness sessions or at training sessions to find out if participants remember accurately what has been discussed. The same tests or parallel forms can be given before and after a session or event so you can compare pre and posttest scores. Or a test can be given at the end of a session so you can decide whether individuals or the group have met a previously agreed on performance standard.

Tests of skill acquisition can also be used to evaluate training sessions to find out if participants have attained a desirable level of proficiency in areas covered by the training session or program.

Tests of skill can take three formats. The first is written simulation. In this format, you present a hypothetical situation and ask respondents to indicate what they would do in that situation. It should be set up so that respondents either select from among answers you provide or write out their answers. For example, you might present trainees with a problem and ask them what to do. You could either provide choices and ask the respondents to select the best one, or you could ask the respondents to write a brief essay.

The second format is role-playing demonstrations. In this format you present a situation and assign roles. The role-playing can be used for the assessment of training. You must take care to establish a non-threatening environment in which individuals feel safe if they make mistakes.

The third format is the production of samples. When using this format you must make sure that the instructions are sufficiently detailed to allow individuals or groups to carry out the task. The criteria for assessing the product must be established beforehand. When establishing the criteria you must ask yourself questions such as: what must be included in the product to make it acceptable? how detailed must it be? how will you indicate that the product is acceptable or not?

Projects often need to develop brief knowledge tests tightly cross-referenced to a curriculum or training program. The following four steps provide a useful summary for the development of a knowledge test.

- Clearly define what you want to learn from the instrument. If you are providing a group of teachers and counselors with training aimed at increasing their awareness and knowledge of nontraditional occupations, then your knowledge test focuses on the essential information you want to communicate. The test items will test whether the essential pieces of information have been understood.

- Link the tests to the evaluation questions. Tests that are being used to detect the impact of a program should be accurate, reliable, and valid.
- Allow time for instrument construction. Test development—no matter how short tests are—takes longer than you anticipate. It takes time to figure out format and then develop items that get to the information you want. Revising and refining each item takes time because the more people read the items, the better they become. Allow time for several cycles of revision.
- Pilot test. Instruments should be tried out by individuals similar to those who will ultimately be respondents.

## STEP #7: List Evaluation Activities and Establish a Timeline for the Various Tasks

When you get to this point, you have clear goals and clearly stated objectives; you have already decided how comprehensive your evaluation is going to be based on a realistic assessment of your resources; you have stated the evaluation questions you want answered; and you have decided what instruments will be used to gather the information you need. The rest is easy. Based on your evaluation question(s), you will make an evaluation plan stating **when, how often,** and **with which groups** evaluation activities will occur.

We used the chart on pages 47 & 48 to help us plan our evaluation activities. The blank charts we provide on pages 49 & 50 can help you to organize your evaluation activities.

## Evaluation Checklist

| | WHO | DATA SOURCE | COMPLETED BY |
|---|---|---|---|
| Revise goals and objectives | Evaluator with Project Director and Project Staff | • Proposal<br>• Project Director<br>• Staff<br>• Other documents | By the end of month 4 |
| Decide on scope of evaluation | Evaluator with Project Director | • Project Director | By the end of month 2 |
| Select/develop evaluation question related to objectives | Evaluator with Project Director | • Project Director<br>• Project Staff<br>• Review of Documents | By the end of month 2 |
| Select/develop instruments To measure objectives | Evaluator with feedback from Project Director | — | By the end of month 3 |
| Review instruments | Evaluator | — | By the end of month 3 |
| Pilot test instruments | Evaluator | — | By the end of month 4 |
| Revise instruments | Evaluator | — | By the end of month 4 |
| Recruit data collectors | Project Director | — | By the end of month 3 |
| Train data collectors | Evaluator | — | By the end of month 4 |
| Select sample | Evaluator | — | By the end of month 3 |
| Copy instruments | Administrative Assistant | — | By the end of month 4 |
| Schedule data collection | Evaluator/ Project Director | — | By the end of month 3 |

## Evaluation Checklist

| | WHO | DATA SOURCE | COMPLETED BY |
|---|---|---|---|
| Administer instruments | Data Collector with assistance from Evaluator | • Students<br>• Teachers<br>• Project Staff<br>• Counselors | By end of month 5 |
| Hire coders | Project Director | ——— | By end of month 4 |
| Train coders | Evaluators | ——— | By end of month 4 |
| Log in data | Coders | ——— | By end of month 6 |
| Summarize data | Evaluator | ——— | By end of month 7 |
| Interpret findings | Evaluator | ——— | By end of month 8 |
| Write evaluation report | Evaluator with feedback from Project Director | Evaluation report will include findings and recommendations | End of evaluation contract |
| Provide feedback to respondents | Project Director | Fact sheets/ press releases, etc. | Before the end of year 1 of project |
| Circulate findings | Project Director | Findings are circulated among all constituencies | " |
| Develop action plan based on findings | Project Director | Action plan is developed with a representative group guided | " |
| Take action | Project Director and Staff | | Second cycle of project activities |

by the project director and/or staff



BEST COPY AVAILABLE

## Evaluation Checklist

| | WHO | DATA SOURCE | COMPLETED BY |
|---|---|---|---|
| Revise goals and objectives | | | |
| Decide on scope of evaluation | | | |
| Select/develop evaluation question | | | |
| Select/develop instruments | | | |
| Review instruments | | | |
| Pilot test instruments | | | |
| Revise instruments | | | |
| Recruit data collectors | | | |
| Train data collectors | | | |
| Select sample | | | |
| Copy instruments | | | |
| Schedule data collection | | | |

## Evaluation Checklist

| | WHO | DATA SOURCE | COMPLETED BY |
|---|---|---|---|
| Administer instruments | | | |
| Hire coders | | | |
| Train coders | | | |
| Log in data | | | |
| Summarize data | | | |
| Interpret findings | | | |
| Write evaluation report | | | |
| Provide feedback to respondents | | | |
| Circulate findings | | | |
| Develop action plan based on findings | | | |
| Take action | | | |

# STEP #8: Estimate the Cost of the Evaluation

The costs of an evaluation include money, time, and energy. The dollar amount can be easily figured out if all the tasks and to do's have been listed and specified and the people involved identified. Once you estimate the number of hours needed to accomplish each task, you can multiply person hours by dollars to come up with a general figure for salary and consultant expenses. Other expenses include travel, supplies, and production costs. Do not underestimate the costs of an evaluation. A "token" evaluation may be inexpensive but will not provide you with a sound knowledge base for program improvement and/or an assessment of your program's effectiveness. It is estimated that the evaluation budget should be between 6% and 10% of the total budget of a program.

We include a chart developed by Bank & Snidman (1981) to estimate the costs of your evaluation.

## Estimating the Costs of the Evaluation

**Personnel**

| Name or Job Title: | # Days | Daily Rate | TOTAL |
|---|---|---|---|
| Project Director | 7 | 150.00 | 1,050 |
| Administrative Assistant | 5 | 75.00 | 375 |
| | | **Staff subtotal:** | 1,425 |

| Consultants: | # Days | Daily Rate | |
|---|---|---|---|
| Evaluator | 30 | 150.00 | 4,500 |
| Test Administrator | 12 | 100.00 | 1,200 |
| Coders | 10 | 48.00 | 480 |
| | | **Consultant subtotal:** | 6,180 |
| | | **Personnel subtotal:** | 7,605 |

**Travel**

| Destination: | # People | # Days | Per Diem Rate | Mileage or Fare | Other Costs | |
|---|---|---|---|---|---|---|
| Boston | 2 | 2 | — | — | — | — |
| Hartford | 2 | 2 | 200.00 | 85.00 | | 285.00 |
| MYC | 2 | 2 | 200.00 | 180.00 | | 380.00 |
| | | | | | **Travel subtotal:** | 665.00 |

**Supplies & Expenses**

| | Amt. Needed | X | Unit Cost | |
|---|---|---|---|---|
| Paper | | | | 50.00 |
| Postage | | | | 25.00 |
| Printing or Copying | | | | 50.00 |
| Computer Analysis | in house | | | — |
| Special Equipment | | | | - |
| Telephone | | | | 25.00 |
| Other | | | | |
| | | | **Supplies & Expenses subtotal:** | 150.00 |
| | | | **TOTAL OF DIRECT EXPENSES:** | 8,420 |

# Estimating the Costs of the Evaluation

**Personnel**

| Name or Job Title: | # Days | Daily Rate | TOTAL |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | Staff subtotal: | |

| Consultants: | # Days | Daily Rate | |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | Consultant subtotal: | |
| | | **Personnel subtotal:** | |

**Travel**

| Destination: | # People | # Days | Per Diem Rate | Mileage or Fare | Other Costs | |
|---|---|---|---|---|---|---|
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | **Travel subtotal:** | |

**Supplies & Expenses**

| | Amt. Needed | X | Unit Cost | |
|---|---|---|---|---|
| Paper | | | | |
| Postage | | | | |
| Printing or Copying | | | | |
| Computer Analysis | | | | |
| Special Equipment | | | | |
| Telephone | | | | |
| Other | | | | |
| | | | **Supplies & Expenses subtotal:** | |
| | | | **TOTAL OF DIRECT EXPENSES:** | |

# STEP #9: Collect, Analyze, and Interpret the Data

Here is one of the descriptive charts in our evaluation report:

TABLE VII

NATIONALITY OF RESPONDENTS

| NATIONALITY | # OF RESPONDENTS | % OF RESPONDENTS |
|---|---|---|
| Puerto Rican | 78 | 45 |
| Dominican | 71 | 40 |
| Cuban | 7 | 4 |
| Central American | 9 | 5 |
| South American | 11 | 6 |
| TOTAL | 176 | 100 |

Now that you are ready to start collecting data, answers to some questions about data collection may help:

**When should you collect data?** The data collection times should be as typical as possible. Not the day before the Christmas holidays, not after some major event has upset the community.

**How can you smooth the wrinkles in the data collection process?** First, tell respondents the purpose of the data collection and be attuned to their feelings. People are more willing to give information when they have been personally approached by an interested individual than when they receive a memo or do not know why the information is needed. Try to minimize the inconveniences of the paperwork for respondents. If respondents get irritated by the methodology, their feelings may color their perception of the events and influence their responses.

**How often should you collect data?** Before answering this question, read your evalua-

Table X is another chart included in our evaluation report. We also include the interpretation of the data.

TABLE X

STATEMENTS RELATED TO MARRIAGE, SCHOOL, WORK, AND WOMEN'S ROLES

| | GIRLS % | | | BOYS % | | |
|---|---|---|---|---|---|---|
| STATEMENT | Yes | Maybe | No | Yes | Maybe | No |
| I think that a woman should choose between having a family and a home and having a job. | 52 | 30 | 17 | 37 | 47 | 16 |
| There is only one good reason for a woman to work outside the home: the family needs the money. | 43 | 35 | 22 | 21 | 37 | 42 |

There seems to be a great deal of uncertainty regarding this issue. More girls than boys seem to think that a woman cannot have a family/house *and* a job.

Forty-three percent of the girls see themselves working *only* if the extra money is needed. These data suggest that the different aspirations (job and career) and expectations of girls and boys should be explored in more depth and that girls should be made aware of alternatives open to them and of how marriage and work roles are being combined by women. Materials to be developed should discuss these issues and make clear that a women does not necessarily have to choose between the two.

tion question(s). The wording of your evaluation questions will determine the number of data collection points. Sometimes you may want to collect information twice—before and after certain program activities; sometimes you may want to collect information on several occasions—when you want to track change over time.

**How many groups should be sources of data?** That depends on the type and number of clients your program serves. You may collect information on one sample of your target population or on the whole group. You may wish to have several groups on your sample—different types of respondents or the same group on various occasions.

The data you collect and the procedures for data collection must reflect the methods of analysis you have selected. These two steps are interrelated. You do not want to collect a handful of data and then discover that the data are not right for the methods of analysis you have selected. Even if you have been conducting an evaluation on your own, at this point you should consider using an evaluation expert because data analysis and interpretation are better left in the hands of someone who has the background and expertise. There are several tasks to be done after the data have been collected. They are:

1. *Summarize the data:* Count and compute—a set of questionnaires, observation sheets—establish percentages and means, look for common characteristics, etc. The results are called descriptive statistics. *Descriptive statistics* allow you to describe your population, its characteristics, or its performance.
2. *Compare sets of data:* When you want to compare sets of data to find out if the difference between them is significant, you will need to use *inferential statistics* — t-tests, analysis of variance (ANOVA), analysis of covariance (ANCOVA), chi-square, etc. These tell you the probability that the observed changes did not occur by chance but were perhaps due to your program. The chart that is on the next page (Davis and Humphreys, 1985) lists some methods of analysis and the occasions where they should be used.
3. *Interpret data:* After you have *analyzed* and *displayed* the data, you must *interpret* the results. The data by themselves will not tell you why things are the way they are. Making inferences from the data means asking yourself "Why is this the way it is?" and "Is this acceptable?" In order to answer these questions you must add data interpretation skills to your knowledge of the area.

The chart on the following page is a summary of various methods for data analysis.

# Key Features of Some Commonly Used Analytical Techniques

(adapted from Owens, 1979)

| | Key Elements | Example |
|---|---|---|
| CHI SQUARE | Used to test the significance of differences between: (1) observed frequencies and expected frequencies; or (2) two sets of observed frequencies. | Is there a significantly larger proportion of girls than boys who applied for entry into a new science program? |
| T-TEST | Used to test whether the differences between two means (group averages) is significant. It can measure the difference between two groups or between the pre- and posttest scores of the same group. | Did the students in the math course show significant growth on the achievement test between the beginning and end of the semester? Did the students in the program score significantly higher on the tests than students not involved in the program? |
| ANALYSIS OF VARIANCE (ANOVA) | Used to test the significance of differences among more than two variables or more than two groups. | Which of three methods of teaching math is most effective for slow learners? |
| MULTIPLE ANALYSIS OF VARIANCE (MANOVA) | Used to test the significance of differences between two or more groups simultaneously on more than one variable. | Do students in an experimental high school program score significantly higher in basic skills (as measured by separate scores in reading, arithmetic, and language expression) than those in the traditional program? |
| ANALYSIS OF COVARIANCE (ANCOVA) | Used to test the significance of differences between two groups when the two groups are not considered equal at pretest time. | Do the students in the work experience program make greater gains in career maturity than those not in the program when students' initial ability is accounted for? |
| MULTIPLE REGRESSION | Combination of several predictive measures to predict achievement or other outcomes. | Do students' age, gender, and pretest scores on an achievement test significantly predict the performance outcomes of the program as measured by posttest scores? |

# STEP #10: Communicate and Act on the Evaluation Findings (Evaluation Report)

You can communicate the findings of your project to many different audiences. If your project is a funded grant, you are often required to write an evaluation report for the grant agency. Secondly, the evaluation report will help with program revision. If your evaluation is also formative—and we think it should be—the evaluation activities and findings will provide the necessary feedback to improve your continuing program, or to provide you with data and conclusions for the next steps or next projects. A third purpose of the final report of your findings is to tell others about your project so that they and you may benefit. This topic will be covered in Section 4.

STEP #10, then, is comprised of two tasks: writing an evaluation report and taking action based on the evaluation findings. An evaluation report should contain the sections listed on the worksheet presented on the next page:

- an explanation of the evaluation;
- a description of the program being evaluated;
- a description of the evaluation procedures including all the instruments, a description of the sample, and rationale for the utilization of the various procedures;
- presentation and discussion of the findings in narrative, tabular, or graphic form;
- conclusions and recommendations based on the findings; and
- a brief summary of what you want to say for those who want to review quickly what you have done. The summary, which is the last piece of the report to be written, contains a summary of evaluation procedures, findings, and recommendations.

## Seven Points to Remember in Writing a Final Report

1. *Think of **who** your audience is*, and focus the report towards their concerns and interests.
2. *Report findings in a useful way*, mindful of others doing similar work.
3. Think of other *possible uses for a final report*—a journal article, a panel at a conference, or as proof for your next proposal.
4. When reporting on your evaluation, be sure to *include* in the Appendix *a copy of the evaluation instruments used.* Samples of instruments are very helpful to readers. The answers received in response to your evaluation instruments should be displayed (graph, tables) or described in the report.
5. At the beginning of the report, *describe your evaluation plan*, including its components, and its relationship to the objectives of the project.
6. *Spell out your project activities* in as much detail as possible. Include such details as numbers and types of people who were involved, number and types of activities, etc.
7. *List each objective* (usually as stated in the proposal) and either provide data to show that each was achieved or discuss why it was not achieved.

Key questions for a Final Evaluation Report are:

What happened?

How have participants benefited?

What have you learned?

Of what use is your learning to someone else?

Errors or failures may be as important as successes in answering this question.

The worksheet on the following page will help you organize your information into a useful evaluation report.

# Components of a Powerful Evaluation Report: A Worksheet

| # | Components | Your Notes/Ideas |
|---|---|---|
| 1. | *Introduction:* Describe the program briefly and summarize the findings of the evaluation. | |
| 2. | *Explanation of the Evaluation Design:*<br>• Outline program goals and objectives.<br>• State the questions the evaluation answers.<br>• Describe the samples you are analyzing, and the sampling techniques you used. Note any problems and limitations.<br>• Explain your information collection methods and the instruments you used. Explain any data collection problems or limitations.<br>• Create visual models and aids to help interpret your findings. | |
| 3. | *Report on the Data:* Explain analytical methods you used and why they are appropriate. | |

| # | Components | Your Notes/Ideas |
|---|---|---|
| 4. | *Presentation of Findings:*<br>• Report all statistical findings.<br>• Display statistics graphically.<br>• Interpret the statistics in terms of the variables and essential issues of the program. | |
| 5. | *Recommendations:*<br>Summarize findings and their significance. Make recommendations for the program. Set forth implications for future development by the program. | |
| 6. | *Appendix:* Assemble credentials of the evaluator and other professionals. Provide essential administrative information (schedules, staffing information). Attach instruments or other documents which may clarify or support the evaluation. | |

While such a "bare bones" final report could be 3-4 pages, a good final report will probably be longer because it will encompass a number of questions and answers about the project's activities. From the details of the answers to those questions, a reader will be able to look at this report and easily understand:

- what you planned to do and why;
- how you planned to evaluate these activities;
- what your outcomes were projected to be;
- whether these occurred; and
- what instruments and other means you used to carry this out.

The following pages provide a Model Final Report that was developed using a 1986 federally funded project.

# Model of a Final Report*

INTRODUCTION (one page)

Describes the project and how the evaluation report is organized.

Example: The purpose of the project *Title IX in Action* was to define a process for the identification of exemplary practices which promoted an equity culture within the elementary school and classroom. Through the collection, codification, and dissemination of equity practices, the project offered a comprehensive approach for extending access to educational opportunities for minority and non-minority females.

SECTION A. Comparison of Actual Performance with Planned Operation

**A. l. Accomplishment of Objectives**

List the objectives of the project (taken from objective section of the original proposal) and the activities conducted to accomplish each.

Example: Objective 1: "to identify the criteria for achieving an equity culture within the school and classroom environment"

Two major activities were designed to accomplish this objective: 1) the development of equity criteria and learner outcomes by expert consultants and 2) the analysis of criteria by practitioners who were committed to and knowledgeable of equity goals and outcomes. (See Attachment 1 for Project Timeline)

**A. 2. Departure from Objectives**

Brief paragraph to indicate any places where objectives were not carried out.

Example: Some slippage in time lines occurred because of school schedules and availability of personnel to participate in the project.

* Taken from: "Title IX in Action," WEEA Grant sponsored by Mid-Continent Regional Education Laboratory, Gretchen Wilbur, Director

SECTION B. Evaluation of Project Operation Outcome

**B.1. Planned Strategies and Methods of Evaluation**

State whether and how evaluation activities occurred throughout the project, and list evaluation procedures used.

Example:
```
Evaluation activities occurred throughout the life
of the project. The evaluation process had the fol-
lowing components:

1. pre-test and post-test evaluation of impact
2. training evaluation, etc.
3. on-site interviews
4. assessment meeting
```

Then each evaluation component was described by a paragraph that told how it was used, reasons for use, people who used it.

Example:
```
1. Pre-test and Post-test

A pre-test and post-test evaluation instrument was
developed for the dual purpose of assessing train-
ing needs and measuring training impact.
```

**B.2. Appropriateness and Effectiveness of the Evaluation**

Discuss how useful these methods were.

Example:
```
By utilizing the four major methods of evaluation
the outcomes of the project were substantively
evaluated. The assessments allowed for cross-checks
and coming at the outcomes from different angles.
```

**B.3. Findings and/or Results of Evaluation**

Describe the findings that resulted from the use of the evaluation procedure.

Example:
```
Findings from comparing the pre-test and post-test
data indicated that the training had an impact on
how the participants addressed equity issues.
```

**B.4. Recommendations**

If you have any recommendations for improvement or if you want to make recommendations for future grantees conducting a similar program, add this section.

**ATTACHMENTS**

All items/lists/documents described in the report should be attached.

Example from model

1. timeline for the project
2. elements of an equity culture
3. training design components
4. some school activities on equity culture that occurred
5. form for cadre member profile
6. evaluation instrument used at training
7. student perception questionnaire

# SECTION 4

## using the evaluation findings

Once you have a complete evaluation report and you know the impact of your program, you may want to disseminate your findings. The worksheet that follows provides help in determining the audience for your evaluation report.

### Determining the Audience for Your Evaluation Report: A Worksheet

| *Who to Disseminate Findings To* | *Your Notes* |
|---|---|
| 1. Other related agencies or programs | |
| 2. All people who provided information for the evaluation | |
| 3. Members of Advisory Board | |
| 4. Potential Sponsors | |
| 5. Funders | |
| 6. Program Staff | |
| 7. Administrators of agencies, institutions, organizations which cooperated in implementing the program | |

## Types of Evaluation Reports*

Give careful consideration to the format of your report and the process you will use to disseminate its information. Below we describe the different types of reports and some common methods for information dissemination.

**Comprehensive Technical Report**

A technical report is a comprehensive description of the evaluation procedures, the program itself, and the evaluation findings with concluding interpretations and recommendations. The technical report is intended as a back-up or source document and should have limited distribution. It would be appropriate for the funding sponsor and the key decision makers who are closely involved in the program.

**Technical Report Supplement**

An appendix or supplement can be added to the copies of the technical report intended for the program staff. This supplement (for in-house review) would contain in-depth information about specific aspects of the program that need to be improved. This supplement might also contain information that is somewhat sensitive or so highly detailed that it would be of interest only to the program staff.

**Executive Summary**

Most audiences receiving an evaluation report are more interested in the results than in the methodology. In addition, most audiences cannot afford the time to read carefully a well-documented and lengthy evaluation report. The executive summary is designed to present a brief overview of the procedures, the important findings, and the general recommendations. Obviously, all the materials presented in the executive summary are contained in the technical report. This summary, usually one to ten pages in length, should make reference to the availability of the technical report. The reader of the summary can glean the important aspects of the study and, if interested, can pursue more detailed information in the technical report. This technique keeps busy administrators informed and increases the potential for utilization of the evaluation data and results.

**Multiple Short Reports**

If there is more than one audience for the evaluation results, several rather short reports targeted for special audiences may be more feasible than a technical report and an executive summary. Evaluators have found that decision-makers prefer short (1 page) and medium (10 pages) reports over longer reports. Brickell also found that most decision-makers request short reports for themselves and longer ones for their subordinates; however, when their subordinates were queried, they requested short reports for themselves and longer ones for their subordinates, and so on down the hierarchy.

**Popularized Reports**

A popularized report is a medium length (10-30 pages) report written without sophisticated terminology in a newsy, information style. The popularized report provides a brief over-

* From *Program Evaluation: Basic Skills Notebook*, The NETWORK, Inc.

# Shabazz 'Report Card' Good, With Exceptions

By RICHARD HAWS
of the State Journal Staff

Malcolm Shabazz experimental high school was found Monday night to be "fairly effective overall in meeting student needs." But in a report on the three-year-old school presented to the School Board, there were some "questions as to the structure and direction of the program."

The report, the work of a citizen committee of the East Educational Area Advisory Council, spoke of the "impression of disorganization."

**Lack of Involvement**

Most frequently mentioned as a weakness of the school was the "lack of involvement by some students."

"Many parents feared that a large number of students were not attending classes and not participating in the program," the report said. Donald Hafeman, East Area attendance director and in charge of the school at 314 N. Sherman Avenue, praised the report for its objectivity, and noted that it was the first citizen review of the controversial school.

**Notable Features**

The evaluators did praise the small class size and resultant low teacher student ratios which it called "notable features which contribute to effective learning."

"The closeness of teacher-student relationships, facilitated by the school's small size, is a major contributor to the personal and academic growth of students," the report noted.

But the evaluators—ten East parents — rated the professional staff as more committed to the controversial school than to the students. "Probably because of their concern for those students who do not appear to be involved at Shabazz, the parents did not rate student involvement and commitment as favorably as they rated the staff involvement," the parents said.

A weakness in the program noted by parents, students, and teachers alike was the lack of adequate supplies, equipment and library facilities. Hafeman said the report would be carefully evaluated and used as a guide toward making possible improvements in the school.

view of the evaluation procedures and describes the program, especially unique or unusual activities, and the results of the evaluation. The results might be portrayed through anecdotes about the impact of the program on individual students as well as summaries of the major findings, conclusions, and recommendations.

### News Releases

Brief highlights of the evaluation findings can be communicated through local newspapers, newsletters, brochures, T.V. spots, and other media. An example of a news release is presented on the next page.

### Oral Presentation

In some cases, oral presentations may have more impact than written ones. The oral presentation would include:

- a short statement on the purpose of the evaluation;
- highlights of evaluation techniques and results for each evaluation purpose;
- recommendations; and
- examples of strategies and techniques for implementing recommendations.

Presenting evaluation results to a group or to an individual provides the audience with an opportunity to question the results, discuss their implications, and determine what steps should be taken to use the results. Through dialogue about the results, a constructive plan of action can be worked out to improve the program, make decisions, and/or gain increased support.

A group presentation could use several speakers (e.g., a student, a teacher, an administrator, an evaluator, and a project director) to discuss the effectiveness of the program from different points of view. Or various individuals could present the results to constituents other than their own. For example, a teacher could present the results to a teachers' association, a parent to the PTO, a student to the student body, and a superintendent to the school board.

**Audio-Visual Presentation**

Evaluation results can be effectively communicated through audio-visual reports as well as written and oral ones. Pictures, transparencies, slides, slide-tape presentations, films, and video tapes can be used to share highlights or key features of the program in an appealing manner. However, most decision makers prefer the print medium to the audio-visual. Audio-visual materials are most useful in group settings for formal presentations.

## Summary

Finally, there are two questions you should ask yourself after spending time and energy in the evaluation of your program. The answers to these questions will give you a sense of purpose.

1. How can we use our findings to improve our program?
2. How can we use our findings to get more funding?

# SECTION 5

## some common misconceptions about evaluation

Here are several misconceptions that regularly surface at the mention of the word evaluation. Program staff often mistakenly make one of the following observations:

### Misconception #1

**Organized fault finding.** *"Evaluation is an effort to find negative information. It will only expose and emphasize program failures."*

A well-managed evaluation is designed to involve staff actively in defining problems they experience in the day-to-day operation of the program. It can help program staff to solve problems and it guides you in brainstorming strategies to support the forces for improvement. A well-designed evaluation will help you and your staff replace vague anxieties which go unstated and unresolved with clear statements of real problems you can solve together.

### Misconception #2

**A team of outsiders poking around.** *"Outside evaluators will never really understand nor care about the mission of our program."*

If you read the "Consumer Tips on Evaluation Services" (page 74) then you will know that your evaluator's experience with projects similar to yours will help him/her to grasp your project's intentions and processes. Make this clear to your staff. Go out of your way to make sure that the evaluator is accepted by key individuals involved in the evaluation.

## Misconception #3

**Viewed as a frill.** *"Evaluation will deplete funds that are badly needed for program activities."*

Evaluation goes hand in hand with detailed planning. Activities which ensure efficient operation of your program can also gather data you will use to demonstrate its effectiveness. Planning and budgeting for evaluation in advance is a way of making evaluation more cost-effective and more supportive of the program. Intelligent planning can save money in evaluation.

## Misconception #4

**The big brother threat.** *"Evaluation does not do anything for our program. It's just a case of excessive monitoring by the funding agency."*

Evaluation measures your success for your sponsor. After showing that you are pursuing worthwhile goals and objectives, you must also demonstrate that you are ready to measure and prove your success in meeting these goals and objectives to that agency's satisfaction. A well-planned and well-conducted evaluation demonstrates your accountability and credibility. A strong evaluation component is a way of showing your prospective funder that you take the issue of accountability seriously and are mobilized to provide that accountability. This orientation will lend important credibility to your efforts.

## Misconception #5

**The "first we run the program, then we evaluate" syndrome.** *"We're too busy running our program. We'll think about evaluation when we finish program activities."*

Evaluation should not be tacked onto a program as an afterthought. How does evaluation fit into the total picture of your program? Program management and evaluation form a feedback loop. Information you gather and lessons you learn from one area will support your efforts in the next, in a cyclic fashion. As you manage your program, you will gather information about program activities and evaluate their role in meeting your objectives. The evaluation procedures provide crucial information you need for effective management of your program.

## Misconception #6

**The bottom line.** *"It's possible to spend only 1 or 2% of our total budget and come up with a good evaluation."*

Your evaluation should be cost-effective, meaning that you get the most for your money. Most experts recommend—if you want more than a "token" or "canned" evaluation—that you spend between 6% and 10% of your total budget. It is difficult to do even a *very* modest evaluation for under *10 staff days*.

# glossary of evaluation terms

**Attitude:** predisposition to respond in a consistent manner with respect to a given object, idea, or experience, based on one's values.

**Audience for the evaluation:** individuals, whether or not they are the clients, who will read, see, or hear of the evaluation.

**Average:** a general term referring to one of three types of measures: mean, median, or mode. Sometimes the average is used colloquially for the mean.

**Baseline data:** information about the performance or attitudes of a group of people prior to their participation in a program or sequence of instruction.

**Benefits:** project outcomes.

**Bias:** factors tending to produce erroneous results.

**Client:** the individual or groups who commission the evaluation, and to whom the evaluator has technical, legal, and immediate responsibilities.

**Closed-ended questions:** questions to which specific response options are given.

**Comparison groups:** a group of people who do not receive the program or instruction under evaluation, but who are administered the questionnaires and tests. A comparison group provides an estimate of how untreated learners perform; this estimate is useful in interpreting change among the treated learners.

**Confound:** a factor or influence which confuses the interpretation of results.

**Content analysis:** a method by which written material from interviews, observations, or open-ended survey terms is organized and reduced to a manageable amount on the basis of a set of rules. Content analysis can focus on any aspect of the material and is reliable to the extent that the rules are specific and consistently applied.

**Control group:** a group, determined by random assignment, that does not experience the program or project under evaluation. Using a control group can eliminate alternative explanations to observed outcomes.

**Cost-effective analysis:** the efficacy of a program in achieving given goals measured in relation to program costs.

**Credibility:** the believed validity; applied to an evaluator and the evaluation.

**Criterion-referenced test:** a type of test which provides information about an individual's (or group's) performance relative to certain independently defined standards, rather than relative to the performance of others.

**Data:** information or material gathered during an evaluation. Data can come from a variety of sources including interviews, questionnaires, tests, observations, and records.

**Descriptive statistics:** numerical summaries, computed from a set of data, that describe various features of the data: overall appearance or shape, typical or average scores, the range of scores, etc.

**Distribution:** a way of ordering a set of data that shows how many individuals obtained a particular score or selected a particular response.

**Effectiveness:** a measure of worth or utility.

**Evaluation:** the process of determining the merit or worth or value of something; the product of that process.

**Forced-choice questions:** multiple choice items that require respondents to select one or more of the given choices. Also called closed-ended questions.

**Formative evaluation:** evaluation done for the purpose of improving a program or product that is still in the process of development or implementation. Results typically go to the program staff.

**Frequency:** the number of times an event occurs (e.g., the "frequency" with which respondents select item A).

**Gain scores:** the difference between pretest scores and posttest scores.

**Impressionistic observation:** a relatively free, unstructured method of observation in which the observer records his or her impressions.

**Inferential statistics:** numerical summaries computed from the data subjected to statistical tests to make inferences about observed outcomes.

**Instrument:** a test, questionnaire, interview protocol, observation schedule, or other tool for gathering data.

**Interview protocol:** interviewer guidelines that describe the questions to be asked, the length of the interview, etc.

**Judgment sample:** a way of selecting a sample in which the evaluator's discretion determines who will be included.

**Likert scale:** a rating scale in which respondents indicate the intensity with which they agree or disagree with a statement. Sample item: "I'm afraid of computers." Likert scale response choices would be "strongly disagree," "disagree," "neither agree nor disagree," "agree," "strongly agree."

**Mean:** the average of a set of scores obtained by summing the scores and dividing by the number of scores.

**Median:** the point of a distribution that divides it in half; 50% of the scores fall below this point and 50% are above it.

**Mode:** the score or value that occurs most frequently in a set of scores or measurements.

**Norm-referenced test:** a type of test that provides information on how well an individual (or group) does in comparison to other individuals taking the same test. Describing work as "average" or "best in the class" or "below average" is norm-referenced language; it says nothing about whether the work is really good or bad.

**Norms:** aids or guides in the interpretation of test results; they describe the typical or average performance of specified subgroups. Norms are not to be regarded as standards for each learner to meet or as desirable levels of attainment.

**Observation schedule:** observer guidelines that describe what is to be recorded and in what format.

**Open-ended questions/format:** test items that require respondents to write out some answers in their own words or speak their own thoughts. Contrasted with forced-choice questions.

**Participant observation:** a method by which an observer participates in the activities of a group or project to gain access to information otherwise difficult to observe.

**Percentile ranks:** whole-number percentage of students in the reference group equaled or surpassed by the individual. If a person obtains a percentile rank of 70, his/her standing is regarded as equaling or surpassing that of (about) 70 percent of the group on which the test was standardized. Maximum percentile rank is 99, by convention.

**Pilot study:** an activity undertaken prior to the evaluation to try out its procedures, instruments, and measures. The tests and questionnaires are administered to a small sample of individuals with characteristics similar to those in the evaluation. Pilot testing instruments help identify problems in administration, interpretation, and scoring.

**Population:** the total number of people, events, or things fitting specified criteria, e.g., all 7th-grade girls in the local high schools.

**Posttest:** measurements of performance or attitudes made after the program, project, or instruction.

**Pretest:** the measurement made before the program or project begins, to gain baseline data and/or provide diagnostic information about program participants.

**Probes:** types of follow-up questions asked to gain more information during an interview.

**Questionnaire:** a self-report measure in which respondents provide written answers to questions probing their attitudes, beliefs, or opinions.

**Random assignment:** a method of selecting a sample that gives every individual in the population an equal chance of being chosen.

**Range:** the difference between the highest and lowest score or values in a given distribution.

**Rating scale:** a scale in which a respondent expresses an opinion, judgment, or evaluation by placing a check mark in one of five or seven categories. A Likert scale is one example.

**Raw score/raw data:** the actual score on a test, usually the number of right answers.

**Reliability**: the consistency among the results from an instrument or person used to measure something.

**Respondents:** people who answer or "respond" to a questionnaire or interview.

**Sample:** that portion of the population that is studied or evaluated.

**Standard deviation:** a measure of the variability, dispersion, or "spread" of a set of scores.

**Standardized test:** tests with standardized instructions for administration, scoring, and interpretation.

**Statistic:** a numerical summary of the characteristics of data from a sample.

**Statistical significance:** mathematical testing to determine the probability that particular results could have occurred by chance alone.

**Stratified random sampling:** a sample produced by subdividing a population on some key characteristics (such as grade level) and then drawing random samples from each stratum.

**Summative evaluation:** evaluation done for the purpose of reporting on the quality and merit of a project. Summative evaluation is conducted for an external client or audience (not the project staff).

**Survey:** an investigation typically using interviews or questionnaires, of the opinions, attitudes, beliefs, or experiences of a group of people.

**Validity:** a test is valid if it measures what it purports to measure. If valid, it must be reliable, but not vice versa.

**Variability:** changeability or lack of consistency.

*(Adapted from: Davis and Humphreys, 1985)*

# consumer tips on evaluation services

Martha Williams at The NETWORK developed a *Program Evaluation Planning Packet* (1976) for project managers which outlines ten tips for consumers of evaluation consultant services. These are:

1. Solicit recommendations for evaluation consultants from other consumers—project directors, school district administrators, state department of education personnel, or a human resource file.
2. Find an educational evaluator, not simply a person with strong statistical or computer background. The consultant should have experience in evaluating educational projects (both process and product evaluation) and be familiar with design constraints and alternatives.
3. When you have identified candidates, ask them for names of other clients they have served. It's wise to talk with a few others who have worked with the consultant to learn about his or her style, expertise, and ability to work within a schedule.
4. Choose a consultant easily accessible to the project. A consultant who is geographically close to the project site can be available for meetings, on-site data collection and reporting, and other aspects of the formative evaluation process. Also, keeping travel costs down helps to make the evaluation cost-effective.
5. Determine the cost of evaluation in advance, based on what is budgeted or what can be transferred to an evaluation line item. A good rule of thumb established the cost of evaluation between three and eight percent of the total project budget.
6. Negotiate with the consultant for frequent on-site visits to discuss procedures, interim results, and problems which arise.
7. Contract carefully with the consultant. When formulating a contract with an outside evaluation consultant, the following things should be considered: (a) who has title to the data—make certain that the project, not the evaluator, has that title; (b) the exact terms of the evaluation—what is expected of whom, when; (c) the number of days on-site; (d) the itemized budget for the evaluation; and (e) a cancellation clause al-

lowing 30-day notice for cancellation of the contract for both the project and the evaluator. The contract should be in writing, signed by both client and consultant.

8. Be sure the consultant is introduced to and accepted by key individuals involved in the evaluation. If the consultant meets resistance in obtaining data, the evaluation will be weakened.
9. Consider having the consultant conduct inservice sessions for project staff and teachers to acquaint them with the purposes and procedures of evaluation. Be sure staff understand the way in which the results of the evaluation will be used.
10. Establish the evaluation design as a cooperative effort involving the project director, the project staff, and the evaluator. This involvement will increase the usefulness of the study and results, take into account reality factors and limitations, and insure that the consultant's interests and orientation don't overly influence the design.

# references


Alkin, M.C., R. Daillak, and P. White. *Using Evaluations: Does Evaluation Make a Difference?* Beverly Hills, CA: Sage, 1979.

Bank, A. and N. Snidman. *Guidebook for Evaluation Dissemination Activities.* Los Angeles, CA: Center for the Study of Evaluation, UCLA, 1981.

Berdie, D.R. and J.F. Anderson. *Questionnaires: Design and Use.* Metuchen, NJ: Scarecrow Press, 1974.

Brinkerhoff, R.O., D.M. Brethower, T. Hluchyi, and J.R. Nowakowski. *Program Evaluation: A Practitioner's Guide for Trainers and Educators.* Hingham, MA: Kluwer-Boston, 1983.

Campbell, D.T. and T. Stanley. *Experimental and Quasi-Experimental Designs for Research.* Chicago, IL: Rand McNally, 1971.

Datta, L.E. and R. Perkoff (Eds.). *Improving Evaluations.* Beverly Hills, CA: Sage Publications, 1979.

Davis, B.G. and S. Humphreys. *Evaluating Intervention Programs: Applications from Women's Programs in Math and Science.* New York, NY: Teachers College Press, 1985.

*Directory of Evaluation Consultants.* New York, NY: The Foundation Center, 1981.

Evaluation Research Society. *Standards for Program Evaluation.* Potomac, MD: Evaluation Research Society, 1980.

*Evaluation Handbook.* San Francisco, CA: Public Management Institute, 1980.

Fink, A. and J. Kosecoff. *An Evaluation Primer.* Beverly Hills, CA: Sage, 1978.

Fitz-Gibbon, C.T. and L.L. Morris. *How to Calculate Statistics.* Beverly Hills, CA: Sage, 1978.

Fitz-Gibbon, C.T. and L.L. Morris. *How to Design a Program Evaluation.* Beverly Hills, CA: Sage, 1978.

Fowler, F. J., Jr. *Survey Research Methods*. Beverly Hills, CA: Sage, 1984.

Gronlund, N.E. *Measurement and Evaluation in Teaching*. 4th Ed. New York, NY: MacMillan, 1982.

Guba, E.G. and Y.S. Lincoln. *Effective Evaluation*. San Francisco, CA: Josey-Bass, 1981.

Joint Committee. *Standards for Evaluation of Educational Programs, Projects, and Materials*. New York, NY: McGraw Hill, 1981.

Levin, H. "Cost Analysis." In N.L. Smith (Ed.) *New Techniques for Evaluation*. Beverly Hills, CA: Sage, 1981.

Mehrens, W.A. and I.J. Lehmann. *Measurement and Evaluation in Education*. New York, NY: Holt, Rinehart & Winston, 1978.

Morris, L.L. and C.T. Fitz-Gibbon. *How to Measure Achievement*. Beverly Hills, CA: Sage, 1978.

Morris, L.L. and C.T. Fitz-Gibbon. *How to Measure Program Implementation*. Beverly Hills, CA: Sage, 1978.

Murphy, J.T. *Getting the Facts*. Santa Monica, CA: Goodyear Publishing, 1980.

Owens, T. and W. Evans. *Program Evaluation Skills for Busy Administrators*. Portland, OR: Northwest Regional Educational Laboratories, 1977.

Scriven, M. "The Methodology of Research." in American Educational Research Association, *Perspectives of Curriculum Evaluation*. Chicago, IL: McNally, 1967.

Stake, R.E., et al. *Evaluating the Arts in Education: A Responsive Approach*. Columbus, OH: Charles E. Merrill, 1975.

Stufflebeam, D.L. (Ed.). *Educational Evaluation and Decisionmaking*. Itasca, IL: F.E. Peacock, 1971.

*notes*

*notes*